Integrating satellite soil moisture estimates and hydrological model products over Australia

M. Khaki[a,1], A. Zerihun[b], J. Awange[a], M. Gibberd[b], A. Dewan[a]

[a] *School of Earth and Planetary Sciences, Spatial Sciences, Curtin University, Perth, Australia.*

[b] *Centre for Crop and Disease Management, School of Molecular and Life Sciences, Curtin University, Perth, Australia.*

Abstract

Accurate soil moisture monitoring is essential for water resource management and agricultural applications and recently it has undertaken using satellite remote sensing or terrestrial hydrological models' products. While both methods have limitations, e.g., the limited soil depth resolution of space-borne data and data deficiencies in models, data assimilation techniques can provide an alternative approach. Here, we use the recently developed data-driven Kalman-Takens approach to integrate satellite soil moisture products with those of the Australian Water Resources Assessment system Landscape (AWRA-L) model. This is done to constrain the model's soil moisture simulations over Australia with those observed from the Advanced Microwave Scanning Radiometer - Earth Observing System (AMSR-E) and Soil Moisture and Ocean Salinity (SMOS) between 2002 and 2017. The main objective is to investigate the ability of the integration framework to improve AWRA-L simulations of soil moisture. The improved estimates are then used to investigate spatio-temporal soil moisture variations. The results show that the proposed model-satellite data integration approach improves the continental soil moisture estimates by increasing their correlation to independent in-situ measurements ($\sim 10\%$ relative to the non-assimilation estimates).

Keywords: Data assimilation, Data-driven, Hydrology, Kalman-Takens, Satellite soil moisture.

Email address: Mehdi.Khaki@postgrad.curtin.edu.au (M. Khaki)

1. Introduction

Soil moisture has a significant impact on hydrology and is essential for agriculture and the broader ecosystem functioning and productivity (Wyland et al., 1996; Lawless et al., 2008; Doraiswamy et al., 2008; Lakhankar et al., 2009). Water contents in the soil surface and root zone layers are critical for applications such as drought monitoring and understanding soil moisture effects on water cycles (e.g., Jupp et al., 1998; Roderick et al., 2014; Enenkel et al., 2016; Xu et al., 2018). Consequently, knowledge of soil moisture status, in time and space, is important for management of water, soil and vegetation resources including fire risk assessment. However, the utility of soil moisture for such applications is dependent on the accuracy of soil moisture monitoring systems.

Soil moisture monitoring methods include, but are not limited to, ground-based measurements, model outputs and use of remotely sensed products. While ground-based measurements are limited spatially to the location of in-situ stations, satellite remote sensing data and hydrological modeling outputs can provide high spatio-temporal resolution data products with a vast coverage (Khaki et al., 2018a). Several studies have applied and validated remotely sensed soil moisture observations over Australia (e.g., De Jeu et al., 2008; Liu et al., 2009; Draper et al., 2009; Su et al., 2013; Holgate et al., 2016). Although these studies have acknowledged that satellite products deliver high spatio-temporal resolution data for soil moisture monitoring, satellite derived soil moisture are limited vertically to the top few centimeters of land surface (Njoku et al., 2003) and accordingly, do not provide information about the root zone soil layer. On the other hand, hydrological models usually provide information on soil moisture at different soil layers. This is very important since soil moisture within the root zone is also essential for the growth of plants. The usefulness of hydrological model outputs, however, can be degraded due to various factors, e.g., data limitations, imperfect modeling, and uncertainties of model parameters (van Dijk et al., 2011; Vrugt et al., 2013; Khaki et al., 2017a,b). To address these issues, data assimilation approach has been used to improve hydrological model estimates by integrating new observations (e.g., Reichle et al., 2002; Alsdorf et al., 2007; Goncalves et al., 2009; Renzullo et al., 2014; Schumacher et al., 2016; Khaki et al., 2017c, 2018b,c). Previous studies have shown that satellite soil moisture datasets can successfully be used for constraining model estimates through data assimilation (e.g., De Jeu et al., 2008; Renzullo et al., 2014; Leroux et al., 2016; Tian et al., 2017).

The present study employs a recently developed data-driven approach, Kalman-Takens filter (Hamilton et al., 2016), to merge soil moisture components from a model and satellites. This data-driven approach does not require a physical model and can perform comparably to dynamic methods at a fraction of the computational cost (Hamilton et al., 2016). The main aim of this work is to constrain the hydrological model soil moisture outputs using satellite measurements to achieve more accurate model-based soil moisture estimates over Australia. For this purpose, the Australian Water Resources Assessment system Landscape model (AWRA-L; van Dijk, 2010) soil moisture storages are updated using the soil moisture products from the Advanced Microwave Scanning Radiometer - Earth Observing System (AMSR-E) and Soil Moisture and Ocean Salinity (SMOS) for the time period of 2002–2017. Using the improved estimates of the proposed Kalman-Takens approach, spatio-temporal soil moisture variations are investigated within Australia.

In what follows, we first describe the data and model used in Section 2, then present the method in Section 3, and discuss the results in Section 4 before concluding the study in Section 5.

2. Model and data

2.1. AWRA-L

The Australian Water Resources Assessment system Landscape (AWRA-L; van Dijk, 2010) model soil moisture outputs are obtained from the Bureau of Meteorology (BoM; http://www.bom.gov.au/). AWRA-L is a one-dimensional grid-based model, which was first developed in 2008 by the Commonwealth Scientific and Industrial Research Organisation (CSIRO) to simulate Australian landscape water stored in vegetation and soil systems. Each data cell is independent of its neighbors and represents different water compartments (van Dijk, 2010; Renzullo et al., 2014). The model parameters include effective soil parameters, water holding capacity and soil evaporation, relating greenness and groundwater recession, and saturated area to catchment characteristics (van Dijk et al., 2013). AWRA-L soil moisture of top, shallow, and deep-root layers broadly correspond to the to 0.10 m, 0.30 m, and 6–10 m depth, respectively (see also Renzullo et al., 2014; Tian et al., 2017). The collected data, i.e., water storages in the top, shallow, and deep root soil layers, covers the entire Australia for the period 2002–2017 at a 0.05-degree spatial resolution.

2.2. Satellite-derived soil moisture

Satellite-based soil moisture products are sourced from the Advanced Microwave Scanning Radiometer for EOS (AMSR-E) for the period 2003–2011 and from European Space Agency (ESA) Soil Moisture Ocean Salinity (SMOS) Earth Explorer mission for the period 2011-2017. The AMSR-E measures the surface brightness temperature, which is correlated to the surface 0-2 cm soil moisture content (Njoku et al., 2003). SMOS Microwave Imaging Radiometer using Aperture Synthesis (MIRAS) radiometer measures the Earth's surface emitted microwave to map land soil moisture for the 0-5 cm depth. Level 3 CATDS (Centre Aval de Traitement des Donnees SMOS) products (Jacquette et al., 2010) with the same spatial resolution ($0.25°\times0.25°$) of AMSR-E are used. SMOS and AMSR-E are selected from ascending and descending passes subject to their higher agreement to in-situ measurements (see, e.g., De Jeu and Owe, 2003; Draper et al., 2009; Jackson and Bindlish, 2012; Su et al., 2013).

The satellite soil moisture products are mainly used to update model state variabilities rather than its absolute values. An additional step is required to prepare these observations for data assimilation by removing the bias between the model simulations and observations. To this end, cumulative distribution function (CDF) matching (Reichle and Koster, 2004; Drusch et al., 2005) is applied to rescale the AMSR-E and SMOS observations. CDF matching relies on the assumption that the difference between observed soil moisture and that of the model is stationary and guarantees that the statistical distribution of both time series is the same (Draper et al., 2009; Renzullo et al., 2014). It should be noted that the satellite products are used to constrain every 0.05-degree grids of AWRA-L output due to the discrepancy between the model and observation spatial resolution. Details of the datasets used in this study are outlined in Table 1.

Table 1: A summary of the datasets used in this study.

Product	**Platform**	**Reference**
Model's soil moisture outputs	AWRA-L	http://www.bom.gov.au/
Soil moisture	AMSR-E	Njoku et al. (2003)
Soil moisture	SMOS	Draper et al. (2009)
Precipitation	TRMM-3B42	Huffman et al. (2007)
Soil moisture in-situ measurements	OzNet	Smith et al. (2012)
Soil moisture in-situ measurements	CosmOz	http://cosmos.hwr.arizona.edu/Probes/australia.php

2.3. In-situ data

In-situ soil moisture measurements are used to validate the results. These observations are obtained from the OzNet moisture-monitoring network (http://www.oznet.org.au/) in the Murrumbidgee catchment (Smith et al., 2012) and CosmOz (http://cosmos.hwr.arizona.edu/Probes/australia.php). The acquired datasets contain volumetric soil moisture measurements at various soil depths. The anomalies of in-situ measurements are estimated to assess the data assimilation results for each layer. While the measurements for the top 10 cm are used to evaluate the top-layer soil moisture estimates, in-situ measurements of the 0–30 cm and 0–90 cm are compared to summations of the model top, shallow, and a portion of deep-root (correspond to measurement thickness) from the model soil moisture.

3. Methodology

3.1. Data Assimilation

The model soil moisture of top, shallow, deep layers are composed of the model state ($\mathbf{x}_t$ at time t). The filtering process conditions a prior probability density function (PDF) of the state with available observations to compute the posterior PDF based on Bayes' rule (Koch, 2007) in two steps; (1) forecasting the state PDF using a dynamical model and (2) updating the forecast PDF by assimilating observations using Bayes' rule (Khaki et al., 2017a). The unscented Kalman filter (UKF) (Julier and Uhlmann, 1997; Julier et al., 2000; Simon, 2006) is used for data assimilation. The filter generates random variables and propagates them through a nonlinear function using a deterministic sampling approach for producing $2L+1$ sigma points with L being the dimension of the state as,

$$\mathbf{x}_t^0 = \mathbf{x}_t, \tag{1}$$

$$\mathbf{x}_t^i = \mathbf{x}_t + \left(\sqrt{(L+\lambda)\mathbf{P}_t}\right)_i \quad i = 1, \ldots, L, \tag{2}$$

$$\mathbf{x}_t^{i+L} = \mathbf{x}_t - \left(\sqrt{(L+\lambda)\mathbf{P}_t}\right)_i \quad i = 1, \ldots, L, \tag{3}$$

where $\mathbf{P}_t$ represents the state covariance matrix. The associated weights to the sigma points are estimated by,

$$w_s^0 = \frac{\lambda}{(L+\lambda)}, \tag{4}$$

$$w_c^0 = \frac{\lambda}{(L+\lambda)} + (1-\alpha^2+\beta), \tag{5}$$

$$w_s^i = \frac{1}{2(L+\lambda)} \quad i = 1,\ldots,2L. \tag{6}$$

λ is the scaling parameter with $\lambda = \alpha^2(L+\beta) - L$. α (0–1) controls the spread of the sigma points and beta is usually set to 0 (Julier and Uhlmann, 1997; Wan and van der Merwe, 2001). The generated sigma points are then integrated with model one time step to calculate the forecast state $\mathbf{x}_{t+1}^f$. To this end, the proposed Kalman-Takens scheme (Hamilton et al., 2016) is used. Contrary to a standard data assimilation filtering, the Kalman-Takens filter does not rely on the physical model (i.e., the model's equations), which significantly decreases computational burden with comparable outcomes.

The Kalman-Takens filter takes advantage of delay-coordinate to replace the dynamical model by a surrogate for advancing the state vector forward in time. The training data, i.e., AWRA-L soil moisture, is used to create this delay-coordinate vector $[\mathbf{x}_t, \mathbf{x}_{t-1}, \ldots, \mathbf{x}_{t-d}]$ (with d the number of delays). A local proxy model $\tilde{f}$ is then created using the M nearest neighbors of the delay-coordinate vectors $[\mathbf{x}_t^1, \mathbf{x}_{t-1}^1, \ldots, \mathbf{x}_{t-d}^1]$, $[\mathbf{x}_t^2, \mathbf{x}_{t-1}^2, \ldots, \mathbf{x}_{t-d}^2], \ldots, [\mathbf{x}_t^M, \mathbf{x}_{t-1}^M, \ldots, \mathbf{x}_{t-d}^M]$, which are selected from the above training data based on the Euclidean distance. The proxy model $\tilde{f}$ is an average of the nearest neighbors (see more details in Hamilton et al., 2016). The forecast state $\mathbf{x}_{t+1}^f$ is then estimated using the nearest neighbors $\mathbf{x}_{t+1}^1$, $\mathbf{x}_{t+1}^2, \ldots, \mathbf{x}_{t+1}^M$.

Once the forecast state is computed, forecast means and corresponding covariance matrices are

calculated following Wan and van der Merwe (2001) as,

$$\mathbf{x}^f_{t+1} = \sum_{j=0}^{2L} \mathbf{w}_s^j \mathbf{x}^{f,j}_{t+1}, \tag{7}$$

$$\mathbf{y}^f_{t+1} = \sum_{j=0}^{2L} \mathbf{w}_s^j \mathbf{H}\mathbf{x}^{f,j}_{t+1}, \tag{8}$$

$$\mathbf{P}^f_{t+1} = \sum_{j=0}^{2L} \mathbf{w}_c^j \left(\mathbf{x}^{f,j}_{t+1} - \mathbf{x}^f_{t+1}\right)\left(\mathbf{x}^{f,j}_{t+1} - \mathbf{x}^f_{t+1}\right)^T + \mathbf{Q}_t, \tag{9}$$

$$\mathbf{P}_{\mathbf{y}^f_{t+1}} = \sum_{j=0}^{2L} \mathbf{w}_c^j \left(\mathbf{H}\mathbf{x}^{f,j}_{t+1} - \mathbf{y}^f_{t+1}\right)\left(\mathbf{H}\mathbf{x}^{f,j}_{t+1} - \mathbf{y}^f_{t+1}\right)^T + \mathbf{R}_{t+1}, \tag{10}$$

$$\mathbf{P}_{\mathbf{x}^f_{t+1},\mathbf{y}^f_{t+1}} = \sum_{j=0}^{2L} \mathbf{w}_c^j \left(\mathbf{x}^{f,j}_{t+1} - \mathbf{x}^f_{t+1}\right)\left(\mathbf{H}\mathbf{x}^{f,j}_{t+1} - \mathbf{y}^f_{t+1}\right)^T. \tag{11}$$

where $\mathbf{Q}_t$ and $\mathbf{R}_{t+1}$ are covariances of noises associated with the process and observation, respectively, and are considered to be Gaussian. $\mathbf{H}$ is the observation operator that maps the model states into observation space to update model top layer storage estimates using the field capacity value to achieve relative wetness (Renzullo et al., 2014). Next, the analysis step is employed, which updates the forecast state $\mathbf{x}^f_{t+1}$ using incoming observations $\mathbf{y}_{t+1}$ to calculate the analysis state $\mathbf{x}^a_{t+1}$ based on the Kalman update equations (see more details in Khaki et al., 2018d),

$$\mathbf{x}^a_{t+1} = \mathbf{x}^f_{t+1} + \mathbf{K}(\mathbf{y}_{t+1} - \mathbf{y}^f_{t+1}), \tag{12}$$

$$\mathbf{K} = \mathbf{P}_{\mathbf{x}^f_{t+1},\mathbf{y}^f_{t+1}} \mathbf{P}^{-1}_{\mathbf{y}^f_{t+1}}, \tag{13}$$

$$\mathbf{P}^a_{t+1} = \mathbf{P}_{\mathbf{x}^f_{t+1}} - \mathbf{K}\mathbf{P}_{\mathbf{y}^f_{t+1}}\mathbf{K}^T. \tag{14}$$

3.2. Evaluation strategy

Once the soil moisture estimates are updated, these are compared with independent in-situ soil moisture measurements (cf. Section 2.3). To this end, soil moisture variation time series of the top, shallow- and deep-root before (AWRA-L raw soil moisture,) and after (from $\mathbf{x}^a$ in Equation 12) data assimilation are spatially interpolated to the location of in-situ stations. The interpolated pre- and post-assimilation results are then compared with the in-situ soil moisture measurements at different depths. For this purpose, we use different soil moisture layers from in-situ measurements including 0-8cm (compared to the model top soil moisture layer), 0-30cm

(compared to the summation of the model top and shallow soil moisture layers), and 0-90cm (compared to the summation of the model top, shallow, and deep soil moisture layers). Note that considering the difference between W3RA outputs (i.e., column water storage measured in mm) and the OzNet measurements (i.e., volumetric soil moisture) and the fact that converting the model outputs into volumetric units may introduce a bias (Renzullo et al., 2014), only a correlation analysis is considered here. It is worth mentioning that cross-correlation is applied to account for lag differences between the time series. We further undertake a significance test for the correlation coefficients using t-distribution. The estimated t-statistic is used to determine the probability of getting a value as large or larger being a random outcome. Correlations coefficients are declared significant at $p < 0.05$. The calculated p-values for the correlations lie under 5% indicating coefficients are significant.

3.3. Principal Component Analysis (PCA)

Historically, principal component analysis (PCA; Lorenz, 1956; Preisendorfer, 1988) has been used to a great effect to extract pattern and to better visualize spatio-temporal variations in hydroclimatological studies. Frappart et al. (2013) found that PCA modes can better represent spatiotemporal variations in time series compared to the full signals by separating dominant water mass change signals (see also Abelen et al., 2015). For this reason, PCA is applied on the assimilation results (cf. Equation 12) to extract to dominant modes (principal components). In this study, as it will be shown (Section 2.3), it is found that the first and second modes of PCA are the most representative soil moisture signals. A schematic illustration of the assimilation process steps is provided in Figure 1.

4. Results

4.1. Validation

We first analyze the impact of data assimilation to determine how well it reflects satellite soil moisture observations in the model estimates. Figure 2 displays the results for the soil top layer, for which assimilation has the largest impact. The reason for this is that satellite soil moisture observations largely represent the moisture variability of the soil top layers, thus, have the greatest effect on the corresponding model soil layer estimate during the assimilation. In Figure 2, we also

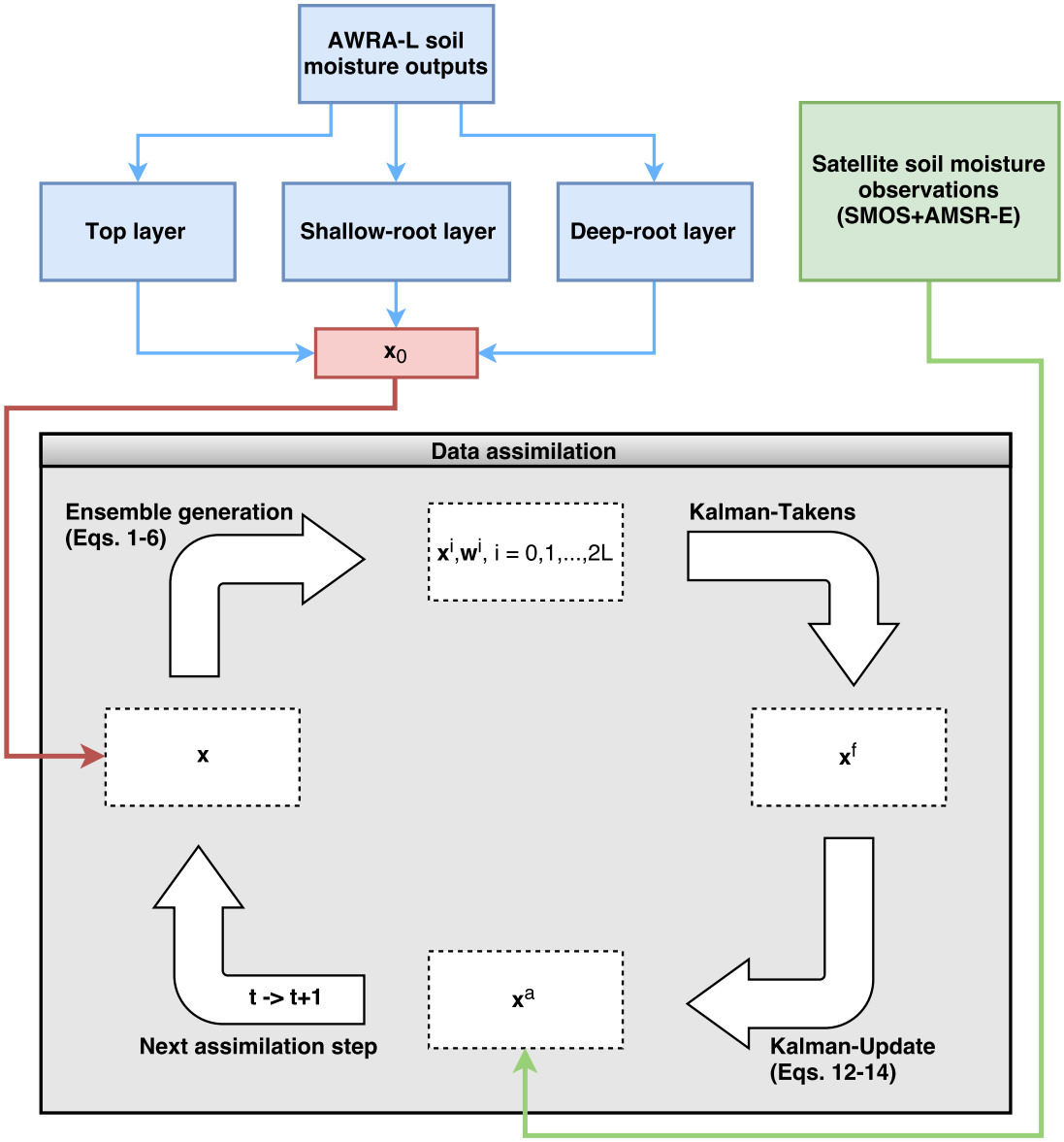

Figure 1: A schematic illustration of the implemented assimilation method. The process starts by reconstructing the dynamics of the system using AWRA-L data. Afterwards, AUKF is used whenever satellite soil moisture observations are available to update model simulations.

plot model-free soil moisture simulations, satellite soil moisture, and rainfall products to evaluate the effect of data assimilation. A misfit is apparent between model-free run estimates and satellite soil moisture observations for the top soil layer (cf. the red and green symbols, Figure 2). Model-data assimilation, however, considerably reduced the mismatch between assimilation estimates and observations (cf. pattern similarity between the green and blue symbols, Figure 2). There is

also larger agreement (∼11%) between assimilation results and precipitation time series compared to that between the precipitation and the model-free run estimates (model outputs without data assimilation). This indicates that data assimilation effectively reflects the climate variabilities in the soil moisture estimates.

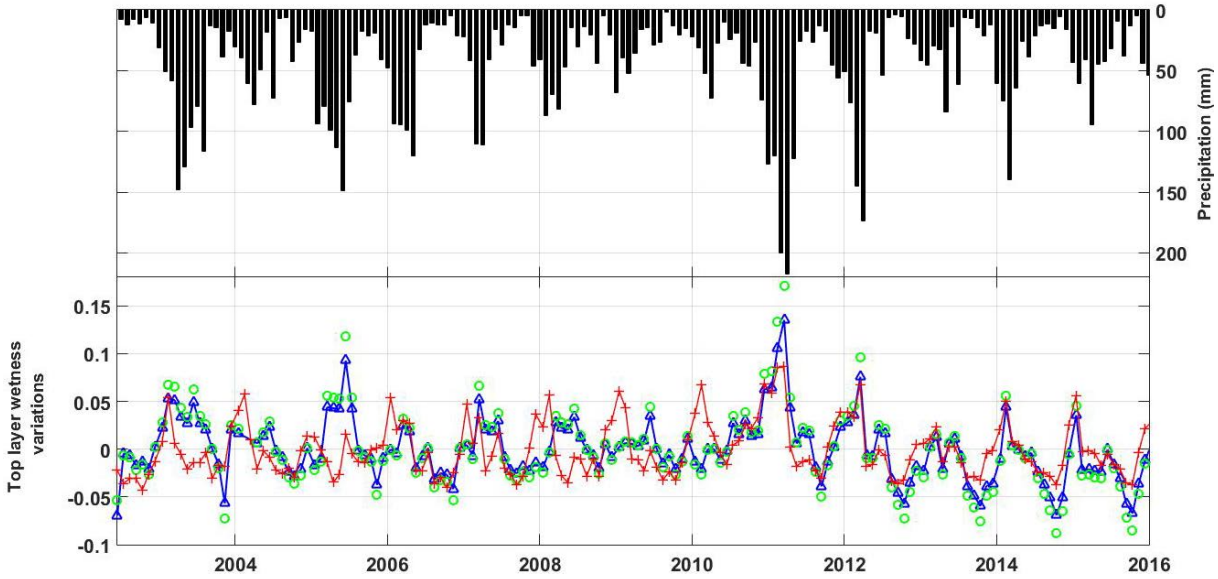

Figure 2: Spatially averaged time series of assimilated top layer soil moisture results (blue), model-free estimates (red), soil moisture observations (green), and precipitation (black) over Australia. The assimilation time series variation pattern can be seen to better match to the precipitation pattern, in particular after 2010.

As mentioned in Section 3.2, to asses the validity of the results, these are compared with independent in-situ measurements. Figure 3 shows the average time series of shallow-root soil moisture for the model-free and assimilation, as well as in-situ observations for all stations. It can be seen that the assimilation reduces misfits between the model simulations and in-situ measurements. This is more evident generally over the large amplitude variations such as those in 2003, 2007, and 2009. In addition, the correlation between both the model-free and assimilation time series and in-situ soil moisture observations are calculated. For each station within the Murrumbidgee catchment, correlations between station observations and the model-free run as well as the data assimilation results are presented in Figure 4. Note that the results are spatially averaged in a 0.3-degree based on in-situ station distances. It is clearly seen that in most of the cases, data assimilation increases the correlation values compared to the model-free run. On average, the magnitude of the correlation coefficients between the post- assimilation soil moisture estimates and the in-situ

199 soil moisture measurements are ∼0.13 units larger than that between model-free run estimates and
200 station observations. This further indicates the capability of data assimilation for improving the
201 model estimates.

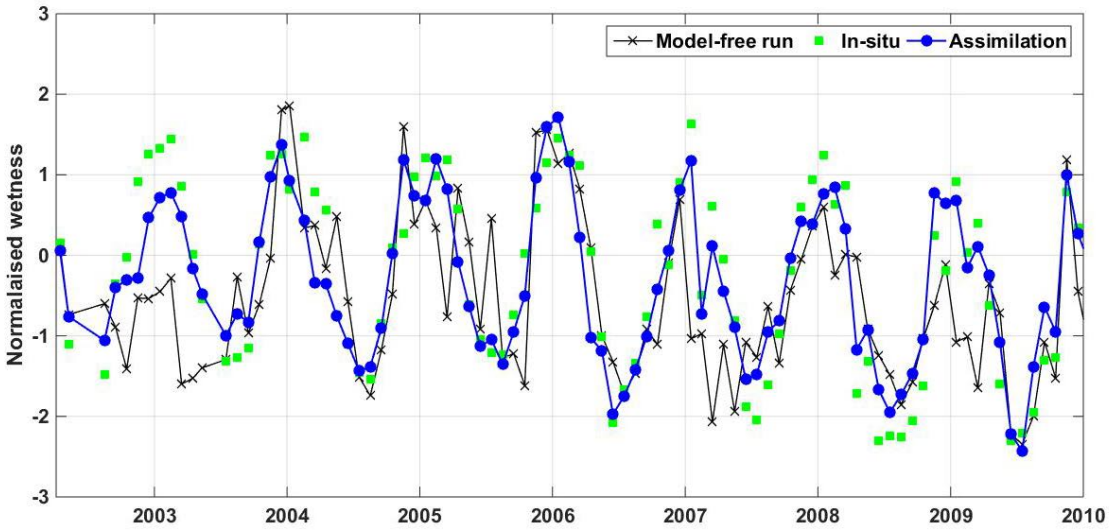

Figure 3: Average soil moisture time series of in-situ measurements (green), model-free run (black), and assimilation (blue) over all in-situ stations. Larger agreement is observed between the assimilation results and those of in-situs.

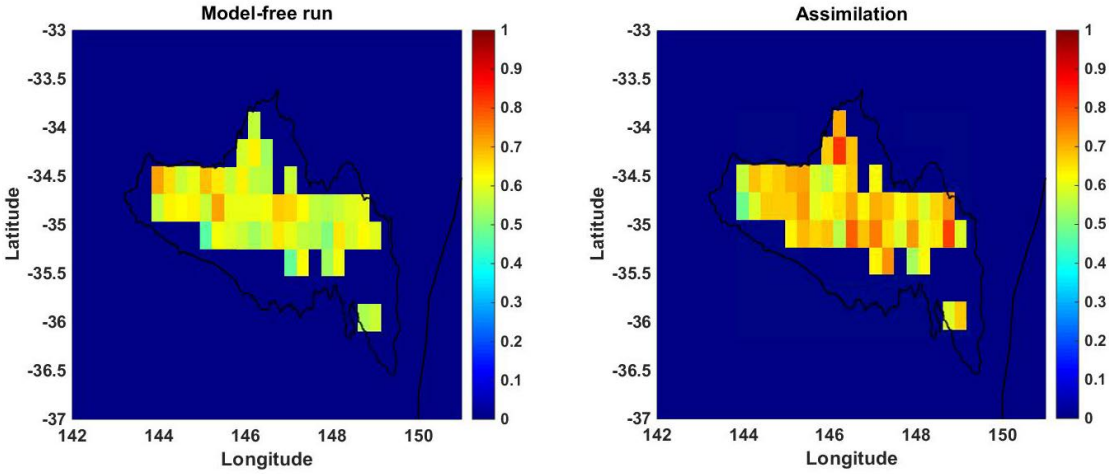

Figure 4: Correlation maps of soil moisture estimates within the Murrumbidgee catchment. The left panel indicates the correlation values between the model-free run results and in-situ measurements, and the right panel shows the correlation values between the assimilation results and in-situ measurements.

202 Detailed validation results of soil moisture estimates against in-situ measurements are presented

in Table 2. The average correlation values of all in-situ stations are reported in this table. It can be seen that data assimilation successfully improves the results of all the soil layers. . Across all soil layers, improved correlations (∼9.69%) are obtained with assimilation results compared to the model-free run. While the highest correlation is achieved for the shallow layer (0.82 correlation value), larger correlation improvements are computed for the top layer (11.74 on average). This indicates the more pronounced impact of soil moisture data assimilation on the top soil layer, which is reflective of the fact that satellite soil moisture observations mostly represent the moisture variability of the soil top layers (2-5 cm). Table 2 shows that integrating satellite-derived soil moisture products with model states, not only improves top layer soil moisture estimates but also positively affects the other layers estimates. This can be seen from the 9.81% and 7.53% correlation improvements for shallow- and deep-root layers, respectively.

Table 2: Summary of the correlations between model-free run estimates and station data, as well as between assimilation estimates and station data for three soil depth layers. For each method, the correlation average values with their 95% confidence interval are presented. Improvements in the assimilation results are calculated as *[(assimilation - model-free run)/model-free run] × 100(%).*

Reference	**< 10 cm**	**< 30 cm**	**< 100 cm**
Model-free run	0.66±0.21	0.75±0.15	0.68±0.18
Assimilation	0.74±0.14	0.82±0.11	0.73±0.12
Average improvement	11.74	9.81	7.53

Further analyses of comparative performance of the data assimilation in estimating of soil moisture are shown in Table 3. Here, the focus is comparison of the magnitude of the correlations of model-free run estimates with satellite soil moisture estimates or precipitation against those of assimilation products with satellite or precipitation data. It can be seen that data assimilation improves correlation with AMSR-E and SMOS measurements (25%) as well as with precipitation (9.7%). These results indicate that data assimilation successfully constrains model simulation with satellite data leading to a better agreement between the results at all layers and assimilated data compared to open loop estimates. There are also considerable improvements in results based on their increased correlation to rainfall variations. These comparisons confirm the evaluation against in-situ measurements and demonstrate that assimilating new observations improves the agreement

between estimates and both in-situ and rainfall measurements by affecting soil moisture variables.

Table 3: Average correlations, as well as correlation improvements of filtered soil moisture estimates with respect to model-free run compared to satellite products and precipitation.

	Reference	< 10 cm	< 30 cm	< 100 cm
Model-free run	AMSR-E+SMOS	0.71	0.66	0.63
	Precipitation	0.82	0.75	0.78
Data assimilation	AMSR-E+SMOS	0.96	0.78	0.76
	Precipitation	0.91	0.82	0.84
Improvements (%)	AMSR-E+SMOS	36.35	17.66	21.08
	Precipitation	11.74	9.81	7.53

4.2. Soil moisture variations

Based on the improved soil moisture estimates, the spatio-temporal variations of the compartments are presented. To begin with, Figure 5 shows the average soil moisture time series for the shallow- and deep-root layers. Compared to the model top layer soil moisture variations in Figure 2, it can be seen that larger variabilities exist in the top, and to a lesser extent in the shallow layers compared to the deep zone soil moisture. Apart from this, similar trends can be observed in all the time series, e.g., a positive trend in 2011 followed by a negative one between 2011 and 2014. The time series for both the shallow and deep layers show strong positive anomalies during 2011, which can be related to the La Niña impact (Forootan et al., 2016). A long-term drought period (2001–2009), known as Millennium Drought (e.g., Ummenhofer et al., 2009; LeBlanc et al., 2012; van Dijk et al., 2013), can also be seen that largely impact both soil moisture estimates, especially the deep-root soil moisture.

Figure 6 shows the temporally averaged soil moisture at the top, shallow- and deep-root layers between 2002 and 2017 at each grid point. Larger soil moisture in all soil layers are available in the northern parts of the country. This could be attributed to the considerable precipitation rate in these areas (see, e.g., Awange et al., 2009, 2011; Forootan et al., 2012, 2016). There are more soil moisture contents over the eastern and southeastern parts. A similar pattern can also be seen

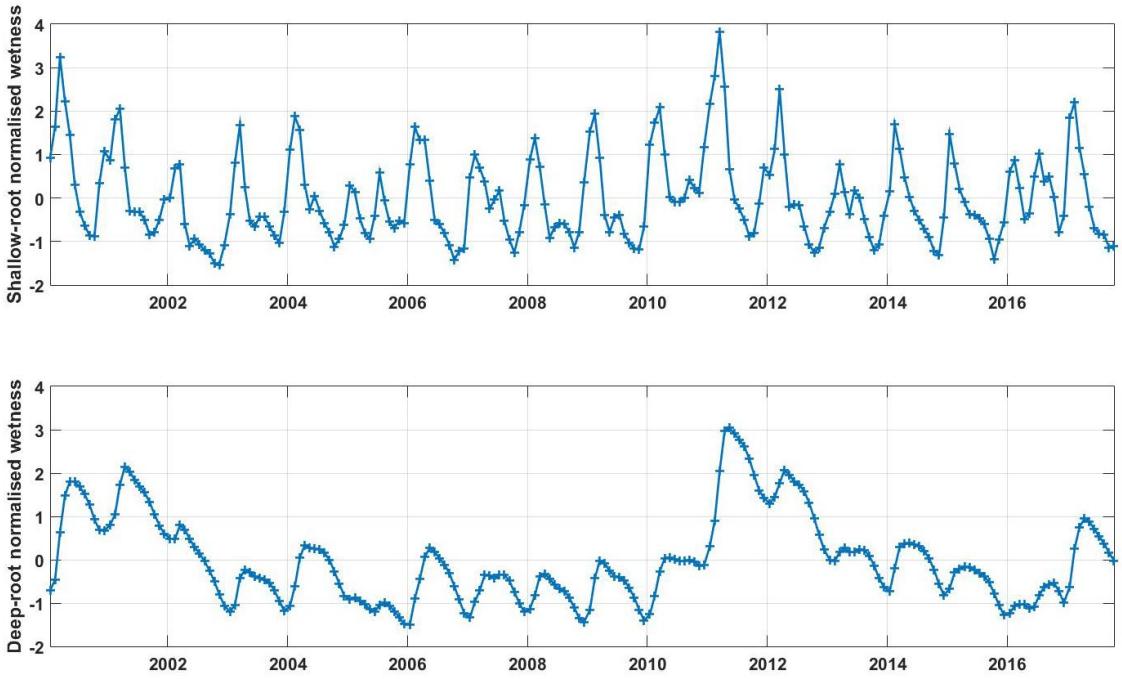

Figure 5: Spatially averaged time series of soil moisture at shallow- and deep-root layers. A negative pattern before 2010 and a significant positive anomaly after 2010 is evident in time series.

in the southwestern parts of Australia. Continentally, the deep root layer and to a lesser degree the shallow root layer show the largest range or variability in soil moisture, where as the top layer soil moisture varies over a narrower range. All the three layers depict large dry regions, in terms of average soil moisture magnitude, in the southern, northwestern and central areas.

Furthermore, average soil moisture maps in January, April, July, and October for the top, shallow- and deep-root layers are displayed in Figure 7. Australia, due to its large size, experiences different (e.g., six) climatic zones (e.g., Fleming et al., 2012) and this translates as two main seasonal patterns, i.e., four seasons in the temperate zone (coastal areas in the southeast and southwest) and a wet/dry pattern in the tropical north areas. It can clearly be seen that larger amounts of soil moisture are available in January and April, especially over the northern part possibly due to the December-February Monsoon rainfall in the north (e.g., Awange et al., 2009, 2011). While smaller changes can be seen for deep-root soil moisture in different months, the top and shallow-root layers demonstrate remarkable variations in various months with larger soil moisture in January and April compared to July and October. In contrast to the northern parts, the eastern part (towards the

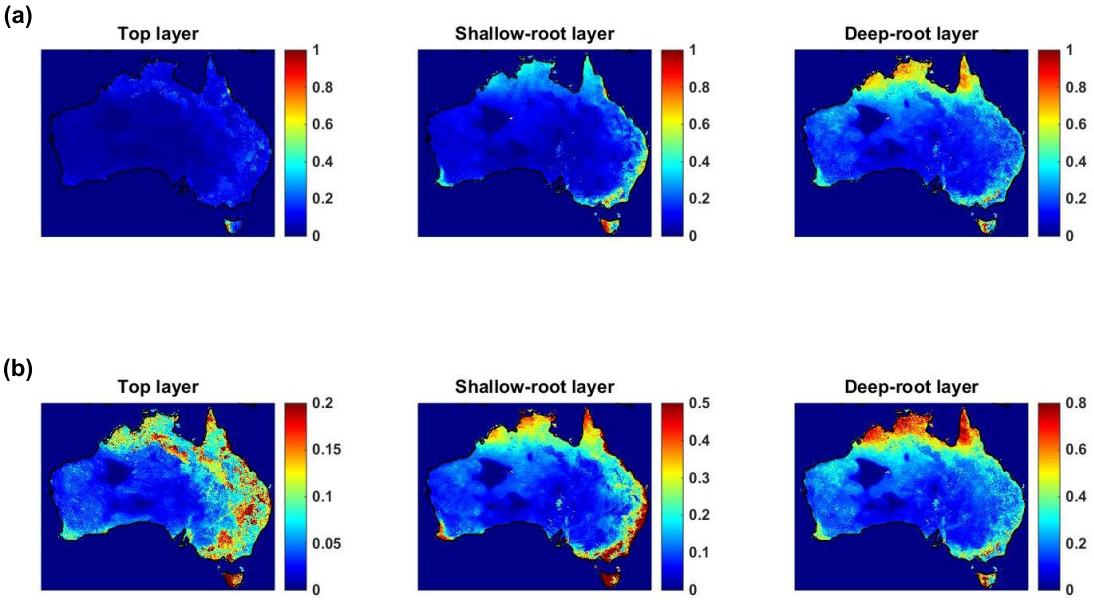

Figure 6: Average (between 2002 and 2017) soil moisture maps (wetness) at different soil layers. The upper row shows values scaled to the same interval [0,1], while the lower bar scales values on a per-dataset basis.

southeast), as well as southwestern parts indicate higher soil moisture amplitudes, especially for the top and shallow-root layers in July and October. In general, the shallow-root soil moisture has a characteristic intra-annual variability. This includes more soil moisture between January and April over northern areas and more soil moisture from July to October over the same areas. The deep-root layer soil moisture (the third column maps on the right hand side) is temporally more stable throughout the year, i.e., the pattern across the seasons is the same continentally contrary to the shallow-root layer where January and April are similar, but differ from July and October. The spatial pattern of soil moisture in all zones follows Figure 6, suggesting that most of the variations can be found in the north and south-west parts.

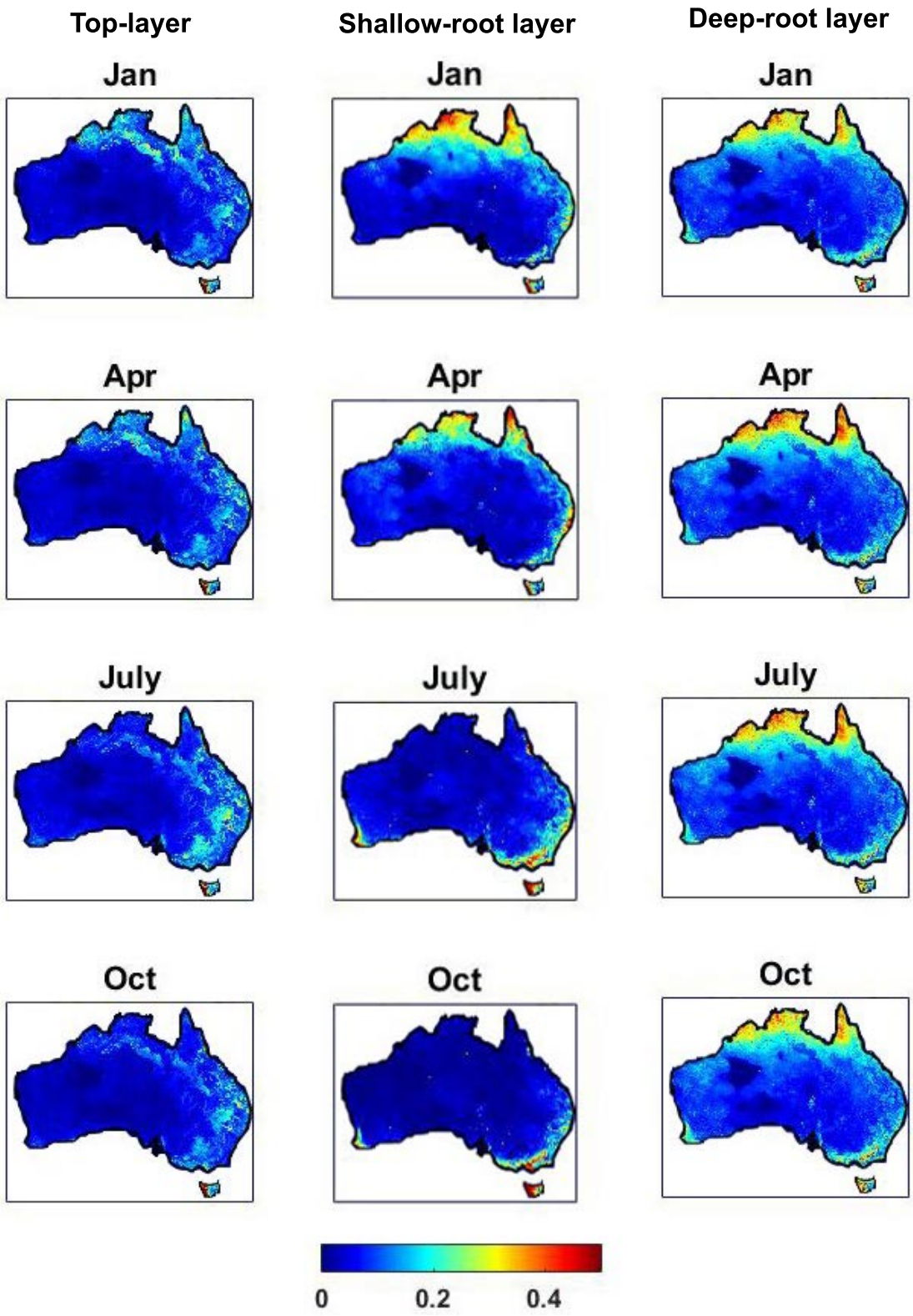

Figure 7: Average soil moisture maps for the months of January, April, July, and October at different soil layers. Note that the top layer maps are multiplied by 2 for a better comparison with the other two layers.

4.3. PCA results

This section presents the PCA results applied on soil moisture estimates from data assimilation. Figure 8 depicts the spatial variations of soil moisture within Australia corresponding to the first two principal components (PCs) of PCA of different soil moisture layers. A similar pattern is evident for all soil moisture layers within the first and second PCs. PC1 maps (accounting for between 43% and 53% of the total variance in soil moisture depending on soil layer) show a distinct north-south continental pattern of variation in soil moisture. By contrast, PC2 (accounting for between 27% and 39% of the remaining total soil moisture variance, depending on soil layer) reveals an east-west variation pattern in soil moisture (Figure 8, panels on the right column). Soil moisture changes in the southwest are captured by the second PC, particularly for the shallow- and deep-root soil moisture layers. The pattern is more regular in the north than other areas. Moreover, patterns in root-zone layers, and especially for the deep-root layer are smoother because these are less sensitive to rain changes. There is generally negative relationship between soil moisture on the north and those in the east, southeast, and southwest, which can be explained by their different climate zones. These soil moisture maps largely follow the soil moisture variation patterns in Figures 6 and 7. Figure 8 suggest that generally higher soil moisture availabilities are apparent in the north, east, and southwest parts of Australia.

The time series corresponding to the first two principal components (PCs) are displayed in Figure 9, which allows us to assess temporal soil moisture variations. Distinct seasonal or intra-annual variabilities can be seen in PC1. PC time series for the top and shallow-root layers show distinct dry and wet patterns in PC1, which is dominant over the north. Positive anomalies exist in 2009 and 2011 from PC1, which is more evident in PC2 for all layers in 2011. This shows considerable soil water increase in a large part of the country due to the pattern of PC1 and PC2, which cover the northern, eastern, and southwestern parts (cf. Figure 8). A remarkable soil moisture increase is observed between 2010 and 2012, which is clearer for the deep-root layer (PC1). PC2 also shows positive strong variations during 2005 and 2016 mostly in the southeastern parts of the continent (see PC2 in Figure 8). Soil moisture PC2 patterns located on the east and northeast are less pronounced mainly due to the impacts of the winter (e.g., Queensland) and summer (e.g., New South Wales) rainfalls. A positive anomaly in deep-root PC1 can be seen between 2009 and 2012, which is followed by a negative trend. The effect of these trends can also be found in PC2 time series, especially between 2010 and 2012 for shallow- and deep-root soil moisture. These large

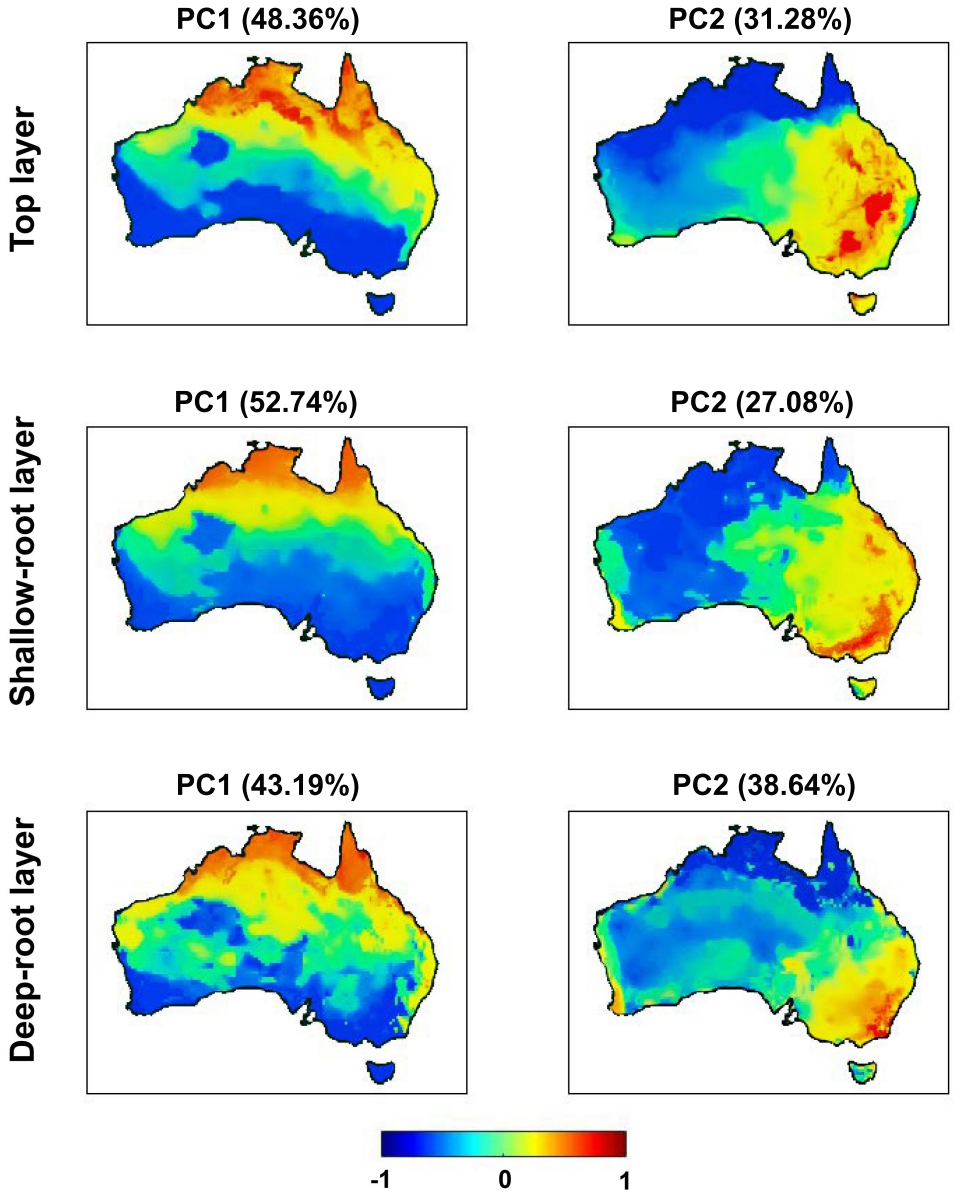

Figure 8: Spatial variations of soil moisture from the first two dominant PCs of PCA for the top, shallow- and deep-root soil moisture layers. Dominant patterns are evident over the north and northeast parts for PC1 and over the east and southeast parts for PC2.

positive anomalies could be due to the effects of the strong La Niña of 2010–2012, which brought above-average rainfall to Australia (Forootan et al., 2016). Several large positive soil moisture amplitudes exist in the time series, e.g., 2004 (PC1), 2008 (PC2), and 2014 (PC2). There are also some negative soil moisture trends in PC time series, e.g., 2003 (PC1), early 2005 (PC2), and 2011 (PC2).

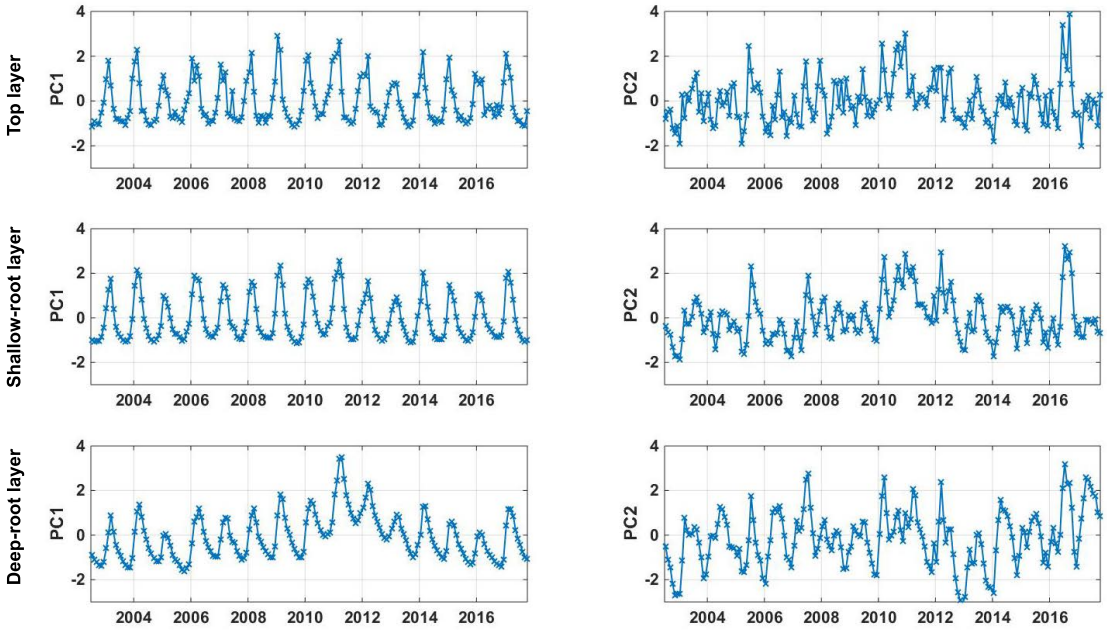

Figure 9: The first two PC time series of soil moisture from PCA for the top, shallow- and deep-root soil moisture layers.

5. Conclusions

The Advanced Microwave Scanning Radiometer - Earth Observing System (AMSR-E) and Soil Moisture and Ocean Salinity (SMOS) soil moisture products are integrated with the Australian Water Resources Assessment system Landscape (AWRA-L) model simulations between 2002 and 2017 over Australia. This is done using a recently developed data-driven approach, Kalman-Takens filter. Based on the evaluation of the assimilation results against independent in-situ soil moisture measurements, we find that the Kalman-Takens filter successfully improves soil moisture estimates. On average, continentally, data-model integration improved soil moisture estimates the correlation between soil moisture estimates and precipitation variations (9.70% on average). The results also show that the data-driven technique can effectively update soil moisture of various layers. Moreover, the assimilation process causes larger improvements in the soil moisture estimates with respect to in-situ measurements in cases where significant trends observed such as the strong La Niña of 2010–2012. These assessments indicate the efficiency of data assimilation for reducing the discrepancies between model simulation results and both in-situ and rainfall measurements. Furthermore, spatio-temporal soil moisture variations of improved soil moisture estimates are investigated. Larger soil

moisture in all soil layers is observed in Australia's north, east, and to a lesser degree in south-west parts. Positive annual trends in soil moisture are found over the western areas.

Acknowledgement

The authors are also grateful to BoM (http://www.bom.gov.au/) for AWRA-L soil moisture simulations. M. Khaki is grateful for the research grant of Curtin International Postgraduate Research Scholarships (CIPRS)/ORD Scholarship provided by Curtin University (Australia).

References

Abelen, S., Seitz, F., Abarca-del-Rio, R., Gntner, A., (2015). Droughts and Floods in the La Plata Basin in Soil Moisture Data and GRACE. Remote Sens., 7, 7324-7349, http://dx.doi.org/10.3390/rs70607324.

Alsdorf, D.E., Rodriguez, E., Lettenmaier, D.P., (2007). Measuring surface water from space, Rev. Geophys., 45, RG2002, http://dx.doi.org/10.1029/2006RG000197.

Awange, J.L., Sharifi, M.A., Baur, O., Keller, W., Featherstone, W.E., Kuhn, M., (2009). GRACE hydrological monitoring of Australia: Current limitations and future prospects. Journal of Spatial Science, 54(1), 2336. https://doi.org/10.1080/14498596.2009.9635164.

Awange, J.L., Fleming, K.M., Kuhn, M., Featherstone, W.E., Heck, B., Anjasmara, I., (2011) On the suitability of the 44 GRACE mascon solutions for remote sensing Australian hydrology, In Remote Sensing of Environment, Volume 115, Issue 3, 2011, Pages 864-875, ISSN 0034-4257, https://doi.org/10.1016/j.rse.2010.11.014.

De Jeu, R.A.M., Owe, M., (2003). Further validation of a new methodology for surface moisture and vegetation optical depth retrieval. Int J Remote Sens 24:45594578, http://dx.doi.org/10.1080/0143116031000095934.

De Jeu, R.A.M., Wagner, W., Holmes, T.R.H., Dolman, A.J., van de Giesen , N.C., Friesen J., (2008) Global Soil Moisture Patterns Observed by Space Borne Microwave Radiometers and Scatterometers, Surveys in Geophysics,Volume 29, Issue 45, pp 399420, http://dx.doi.org/10.1007/s10712-008-9044-0.

342 Doraiswamy, P.C., Hatfield, J.L., Jackson, T.J., Akhmedov, B., Prueger, J., Stern, A., (2008). Crop
343 condition and yield simulations using Landsat and MODIS, Remote Sensing of Environment, 92
344 (2004) 548-559. [6] C. Lawless, M. A. Semenov, and P. D. Jamieson, Quantifying the effect of
345 uncertainty in soil moisture characteristics on plant growth using a crop simulation model, Field
346 Crops Research, 106, 138-147.

347 Draper, C.S., Mahfouf, J.-F., Walker, J.P., (2009), An EKF assimilation of AMSR-
348 E soil moisture into the ISBA land surface scheme, J. Geophys. Res., 114, D20104,
349 http://dx.doi.org/10.1029/2008JD011650.

350 Drusch, M., Wood, E.F., Gao, H., (2005). Observation operators for the direct assimilation of
351 TRMM microwave imager retrieved soil moisture. Geophysical Research Letters, 32, L15403.

352 Enenkel, M., Reimer, C., Dorigo, W., Wagner, W., Pfeil, I., Parinussa, R., De Jeu, R., (2016).
353 Combining satellite observations to develop a global soil moisture product for near-real-time ap-
354 plications, Hydrology and Earth System Sciences, 20:4191-4208, http://dx.doi.org/10.5194/hess-
355 20-4191-2016.

356 Fleming, P.J.S., Allen, B.L., Ballard, G.A., (2012). Seven considerations about dingoes as biodi-
357 versity engineers: the socioecological niches of dogs in Australia. Australian Mammalogy 34,
358 119131, http://dx.doi.org/10.1071/AM11012

359 Forootan, E., Awange, J.L., Kusche, J., Heck, B., Eicker, A., (2012). Independent patterns of water
360 mass anomalies over Australia from satellite data and models, In Remote Sensing of Environment,
361 Volume 124, 2012, Pages 427-443, ISSN 0034-4257, https://doi.org/10.1016/j.rse.2012.05.023.

362 Forootan, E., Khandu, Awange, J.L., Schumacher, M., Anyah, R.O., van Dijk, A.I.J.M., Kusche,
363 J., (2016). Quantifying the impacts of ENSO and IOD on rain gauge and remotely sensed pre-
364 cipitation products over Australia, In Remote Sensing of Environment, Volume 172, 2016, Pages
365 50-66, ISSN 0034-4257, https://doi.org/10.1016/j.rse.2015.10.027.

366 Frappart, F., Seoane, L., Ramillien, G., (2013). Validation of GRACE-derived terrestrial water
367 storage from a regional approach over South America. Remote Sensing of Environment, Elsevier,
368 2013, 137, pp.69-83.

369 Hamilton, F., Berry, T., Sauer, T., (2016). Ensemble Kalman Filtering without a Model, Phys.
370 Rev. X 6, 011021, Vol. 6, Iss. 1, http://dx.doi.org/10.1103/PhysRevX.6.011021.

Goncalves, L.G., Shuttleworth, W.J., Vila, D., Larroza, E., Bottino, M.J., Herdies, D.L., Aravequia, J.A., De Mattos, J.G., Toll, D.L., Rodell, M., Houser, P., (2009). The South American Land Data Assimilation System (SALDAS) 5-Yr Retrospective Atmospheric Forcing Datasets. J. Hydrometeor., 10, 9991010, https://doi.org/10.1175/2009JHM1049.1.

Harris, C.R., (1964). Influence of Soil Type and Soil Moisture on the Toxicity of Insecticides in Soils to Insects, Nature, 202, 724-724.

Holgate, C.M., De Jeu, R.A.M., van Dijk, A.I.J.M, Liu, Y.Y., Renzullo, L.J., Vinodkumar, Dharssi, I., Parinussa, R.M., Van Der Schalie, R., Gevaert, A., Walker, J., McJannet, D., Cleverly, J., Haverd, V., Trudinger, C.M., Briggs, P.R., (2016). Comparison of remotely sensed and modelled soil moisture data sets across Australia, In Remote Sensing of Environment, Volume 186, Pages 479-500, ISSN 0034-4257, https://doi.org/10.1016/j.rse.2016.09.015.

Huffman, G.J., Adler, R.F., Bolvin, D.T., Gu, G., Nelkin, E.J., Bowman, K.P., Hong, Y., Stocker, E.F., Wolff, D.B., (2007). The TRMM Multi-satellite Precipitation Analysis: Quasi- Global, Multi-Year, Combined-Sensor Precipitation Estimates at Fine Scale. J. Hydrometeor., 8(1), 38-55.

Jackson, T., Bindlish, R., (2012). Validation of Soil Moisture And Ocean Salinity (SMOS) soil moisture over watershed networks in the US, IEEE Trans. Geosci. Remote Sens., 50, 15301543.

Jacquette, E., Al Bitar, A., Mialon, A., Kerr, Y., Quesney, A., Cabot, F., et al. (2010). SMOS CATDS level 3 global products over land. In C. M. U. Neale, A. Maltese (Eds.), Remote Sensing for Agriculture, Ecosystems, and Hydrology XII. volume 7824 of Proceedings of SPIE-The International Society for Optical Engineering. Conference on Remote Sensing for Agriculture, Ecosystems, and Hydrology XII, Toulouse, France.

Julier, S.J., Uhlmann, J.K., (1997). A New Extension of the Kalman Filter to Nonlinear Systems. In Proc. of AeroSense: The 11th Int. Symp. on Aerospace/Defence Sensing, Simulation and Controls.

Julier, S., Uhlmann, J., Durrant-Whyte, H.F., (2000). A new method for the nonlinear transformation of means and covariances in filters and estimators, IEEE Trans. Automat. Control 45, 477-482.

Jupp, D.L.B., Guoliang, T., McVicar, T.R., Yi, Q., Fuqin, L., (1998). Soil moisture and drought monitoring using remote sensing. I: Theoretical background and methods. In: CSIRO Earth Observation Centre.

Kamthonkiat, D., Honda, K., Turral, H., Tripathi, N.K., Wuwongse, V., (2005). Discrimination of irrigated and rainfed rice in a tropical agricultural system using SPOT VEGETATION NDVI and rainfall data, International Journal of Remote Sensing, 26, 2527-2547.

Khaki, M., Hoteit, I., Kuhn, M., Awange, J., Forootan, E., van Dijk, A.I.J.M., Schumacher, M., Pattiaratchi, C., (2017a). Assessing sequential data assimilation techniques for integrating GRACE data into a hydrological model, Advances in Water Resources, Volume 107, Pages 301-316, ISSN 0309-1708, http://dx.doi.org/10.1016/j.advwatres.2017.07.001.

Khaki, M., Schumacher, M., J., Forootan, Kuhn, M., Awange, E., van Dijk, A.I.J.M., (2017b). Accounting for Spatial Correlation Errors in the Assimilation of GRACE into Hydrological Models through localization, Advances in Water Resources, Available online 1 August 2017, ISSN 0309-1708, https://doi.org/10.1016/j.advwatres.2017.07.024.

Khaki, M., Ait-El-Fquih, B., Hoteit, I., Forootan, E., Awange, J., Kuhn, M., (2017c). A two-update ensemble Kalman filter for land hydrological data assimilation with an uncertain constraint, In Journal of Hydrology, Volume 555, 2017, Pages 447-462, ISSN 0022-1694, https://doi.org/10.1016/j.jhydrol.2017.10.032.

Khaki, M., Forootan, E., Kuhn, M., Awange, J., Longuevergne, L., Wada, W., (2018a). Efficient Basin Scale Filtering of GRACE Satellite Products, In Remote Sensing of Environment, Volume 204, Pages 76-93, ISSN 0034-4257, https://doi.org/10.1016/j.rse.2017.10.040.

Khaki, M., Forootan, E., Kuhn, M., Awange, J., Papa, F., Shum, C.K., (2018b). A Study of Bangladesh's Sub-surface Water Storages Using Satellite Products and Data Assimilation Scheme, Science of The Total Environment, Volume 625, 2018, Pages 963-977, ISSN 0048-9697, https://doi.org/10.1016/j.scitotenv.2017.12.289.

Khaki, M., Forootan, E., Kuhn, M., Awange, J., van Dijk, A.I.J.M., Schumacher, M., Sharifi, M.A., (2018c). Determining Water Storage Depletion within Iran by Assimilating GRACE data into the W3RA Hydrological Model. Advances in Water Resources, 114:1-18, https://doi.org/10.1016/j.advwatres.2018.02.008.

Khaki, M., Hamilton, F., Forootan, E., Hoteit, I., Awange, J., Kuhn, M., (2018). Non-parametric Data Integration scheme for Land Hydrological Applications. Revised and resubmitted to Water Resource Research.

Koch, K.R., (2007). Introduction to Bayesian Statistics (2nd), Springer.

Lakhankar, T., Krakauer, N., Khanbilvardi, R., (2009). Applications of microwave remote sensing of soil moisture for agricultural applications. Int. J. Terraspace Sci. Eng. 2 (1), 8191.

Lawless, C., Semenov, M.A., Jamieson, P.D., (2008). Quantifying the effect of uncertainty in soil moisture characteristics on plant growth using a crop simulation model, Field Crops Research, 106, 138-147.

LeBlanc, M., Tweed, S., Van Dijk, A., Timbal, B., (2012). A review of historic and future hydrological changes in the Murray Darling Basin. Global Planetary Change (8081): 226246.

Lee, E., Chase, T.N., Rajagopalan, B., Barry, R.G., Biggs, T.W., Lawrence, P.J., (2009). Effects of irrigation and vegetation activity on early Indian summer monsoon variability, International Journal of Climatology, 29, 573-581.

Leroux, D.J., Pellarin, T., Vischel, T., Cohard, J.-M., Gascon, T., Gibon, F., Mialon, A., Galle, S., Peugeot, C., Seguis, L., (2016). Assimilation of SMOS soil moisture into a distributed hydrological model and impacts on the water cycle variables over the Oum catchment in Benin, Hydrol. Earth Syst. Sci., 20, 2827-2840, https://doi.org/10.5194/hess-20-2827-2016, 2016.

Liu, Y.Y., Mccabe, M.F., Evans, J.P., van dijk, A.I.J.M., De Jeu, R.A.M., Su, H., (2009). Comparison of soil moisture in GLDAS model simulations and satellite observations over the Murray Darling Basin [in:] R.S. Anderssen, R.D. Braddock, L.T.H. Newham (eds.), Proceedings of the International Congress on Modelling and Simulation, Cairns, Australia, 13-17 July, pp. 2798-2804.

Lorenz, E., (1956). Empirical orthogonal function and statistical weather prediction. Technical Report Science Report No 1, Statistical Forecasting Project. MIT, Cambridge.

Njoku, E.G. et al. (2003). Soil moisture retreival from AMSR-e. IEEE Transactions on Geo-science and Remote Sensing. 41:2, 215-229.

455 Preisendorfer, R.W., (1988). Principal Component Analysis in Meteorology and Oceanography.
456 Elsevier, New York, 425 pp.

457 Productivity Commission (2005). Trends in Australian Agriculture, Research Paper, Canberra.

458 Reichle, R.H., McLaughlin, D.B., Entekhabi, D., (2002). Hydrologic Data Assimilation with
459 the Ensemble Kalman Filter. Mon. Wea. Rev. 130, 103114, http://dx.doi.org/10.1175/1520-
460 0493(2002)130¡0103:HDAWTE¿2.0.CO;2.

461 Reichle, R.H., Koster, R.D., (2004). Bias reduction in short records of satellite soil moisture, Geo-
462 phys. Res. Lett., 31, L19501, http://dx.doi.org/10.1029/2004GL020938.

463 Renzullo, L.J., Van Dijk, A.I.J.M., Perraud, J.M., Collins, D., Henderson, B., Jin, H.,
464 Smith, A.B., McJannet, D.L., (2014). Continental satellite soil moisture data assimilation im-
465 proves root-zone moisture analysis for water resources assessment. J. Hydrol., 519, 27472762.
466 http://dx.doi.org/10.1016/j.jhydrol.2014.08.008.

467 Roderick, M.L., Sun, F., Lim, W.H., Farquhar, G.D., (2014). A general framework for understand-
468 ing the response of the water cycle to global warming over land and ocean, Hydrol. Earth Syst.
469 Sci., 18, 1575-1589, https://doi.org/10.5194/hess-18-1575-2014.

470 Schumacher, M., Kusche, J., Dll, P., (2016). A systematic impact assessment of GRACE
471 error correlation on data assimilation in hydrological models, Journal of Geodesy,
472 http://dx.doi.org/10.1007/s00190-016-0892-y.

473 Simon, D., (2006). Optimal state estimation: Kalman, H_∞, and nonlinear approaches, John Wiley
474 and Sons.

475 Smith, A.B., Walker, J.P., Western, A.W., Young, R.I., Ellett, K.M., Pipunic, R.C., Richter, H.,
476 (2012). The Murrumbidgee soil moisture monitoring network data set. Water Resour. Res. 48
477 (7), 16. http://dx.doi.org/10.1029/2012WR011976.

478 Su, C.-H., Ryu, D., Young, R.I., Western, A.W., Wagner, W., (2013). Inter-comparison of mi-
479 crowave satellite soil moisture retrievals over the Murrumbidgee Basin, southeast Australia. Re-
480 mote Sensing of Environment, 134, 111.

481 Thenkabail, P.S., Biradar, C.M., Noojipady, P., Dheeravath, V., Li, Y., Velpuri, M., Gumma, M.,
482 Gangalakunta, O.R.P., Turral, H., Cai, X., Vithanage, J., Schull, M.A., Dutta, R., (2009). Global

irrigated area map (GIAM), derived from remote sensing, for the end of the last millennium, International Journal of Remote Sensing, 30, 3679 - 3733.

Tian, S., Tregoning, P., Renzullo, L.J., van Dijk, A.I.J.M., Walker, J.P., Pauwels, V.R.N., Allgeyer, S., (2017). Improved water balance component estimates through joint assimilation of GRACE water storage and SMOS soil moisture retrievals, Water Resour. Res., 53, http://dx.doi.org/10.1002/2016WR019641.

Ummenhofer, C.C., England, M.H., McIntosh, P.C., Meyers, G.A., Pook, M.J., Risbey, J.S., Sen Gupta, A., Taschetto, A.S., (2009). What causes southeast Australia's worst droughts? Geophys. Res. Lett. 36, L04706. http://dx.doi.org/10.1029/2008GL036801.

van Dijk, A.I.J.M., (2010). The Australian Water Resources Assessment System: Technical Report 3, Landscape model (version 0.5) Technical Description, CSIRO: Water for a Healthy Country National Research Flagship.

van Dijk, A.I.J.M., Renzullo, L.J., and Rodell, M., (2011). Use of Gravity Recovery and Climate Experiment terrestrial water storage retrievals to evaluate model estimates by the Australian water resources assessment system, Water Resour. Res., 47, W11524, http://dx.doi.org/10.1029/2011WR010714.

van Dijk, A.I.J.M., Pea-Arancibia, J.L., Wood, E.F., Sheffield, J., Beck, H.E., (2013). Global analysis of seasonal streamflow predictability using an ensemble prediction system and observations from 6192 small catchments worldwide, Water Resour. Res., 49, 27292746, http://dx.doi.org/10.1002/wrcr.20251.

Vrugt, J.A., ter Braak, C.J.F., Diks, C.G.H., Schoups, G., (2013). Advancing hydrologic data assimilation using particle Markov chain Monte Carlo simulation: theory, concepts and applications, Advances in Water Resources, Anniversary Issue - 35 Years, 51, 457-478, http://dx.doi.org/10.1016/j.advwatres.2012.04.002.

Wan, E., van der Merwe, R., (2001). The Unscented Kalman Filter. Wiley Publishing.

Wyland, L.J., Jackson, L.E., Chaney, W.E., Klonsky, K., Koike, S.T., Kimple, B., (1996). Winter cover crops in a vegetable cropping system: Impacts on nitrate leaching, soil water, crop yield, pests and management costs, Agriculture, Ecosystems & Environment, 59, 1-17.

Xu, Y., Wang, L., Ross, K.W., Liu, C., Berry, K., (2018). Standardized Soil Moisture Index for Drought Monitoring Based on Soil Moisture Active Passive Observations and 36 Years of North American Land Data Assimilation System Data: A Case Study in the Southeast United States. Remote Sens., 10(2), 301, http://dx.doi.org/10.3390/rs10020301.