Operational soil moisture data assimilation for improved continental water balance prediction

Siyuan Tian[1]*, Luigi J. Renzullo[1], Robert C. Pipunic[2], Julien Lerat[2], Wendy Sharples[2], Chantal Donnelly[2]

[1] Fenner School of Environment & Society, The Australian National University, ACT, 2601, Australia

[2] Water Program, Bureau of Meteorology, Australia

Corresponding author: Siyuan Tian (siyuan.tian@anu.edu.au)

Key Points:

- We develop a simple and efficient method to improve operational soil water balance model through satellite data assimilation.
- We suggest an approach to use analyzed surface soil moisture estimates to impart mass conservation constraints on related states and fluxes
- The impact of the satellite data assimilation on model estimates of soil moisture can persist for several weeks.

Abstract

A simple and efficient method was developed to improve soil moisture representation in an operational water balance model through satellite data assimilation. The proposed method exploits temporal covariance statistics between modelled and satellite-derived soil moisture to produce analysed estimates, as a weighted combination of all data sources. We demonstrate the application of the method to the Australian Water Resources Assessment (AWRA) model and evaluate the accuracy of the approach against in-situ observations across the water balance. The correlation between simulated surface soil moisture and in-situ observation is increased from 0.54 (open-loop) to 0.77 (data assimilation). We suggest an approach to use analysed surface moisture estimates to impart mass conservation constraints on related states and fluxes of the AWRA model in a post-analysis adjustment. The improvements gained from data assimilation can persist for more than one week in surface soil moisture estimates and one month in root-zone soil moisture estimates.

Plain Language Summary

The access to accurate daily continental soil water balance predictions is valuable for water management practitioners, policy makers and researchers in support of water resources assessment and agriculture planning. This study develops a simple and robust method for an operational water balance model to incorporate satellite soil moisture products for improved accuracy and spatial representation of soil water storage predictions. The integration of satellite soil moisture products can provide persistent constraints in model predictions for up to several weeks.

1 Introduction

Accurate estimation of soil moisture is fundamental to monitoring and forecasting water availability and land surface conditions under extreme events such as droughts, heatwaves and floods (Ines et al., 2013; Sheffield and Wood, 2007; Tian et al., 2019a; Wanders et al., 2013). The assimilation of satellite soil moisture into land surface and hydrology models has been repeatedly demonstrated to improve model representation of soil water dynamics (Draper et al., 2012; Kumar et al., 2009; Pipunic et al., 2008; Reichle and Koster, 2005; Renzullo et al., 2014; Tian et al., 2017; Tian et al., 2019b). Soil moisture is the linchpin between atmospheric fluxes, surface- and ground-water hydrology, thus it is important that any changes in modelled estimates are not detrimental to other components of the water balance.

Soil moisture anomalies can persist for months (Vinnikov et al., 1996), but the spatial pattern can vary significantly due to the heterogeneous spatial distribution of rainfall and variability in soil properties, land cover type and topography. Due to this large spatial variability of soil moisture, the utility of ground-based, point-scale measurements is limited. Soil moisture estimates from land surface models are adversely affected by the uncertainties of atmospheric forcing, model dynamics and model parameterization. Remotely sensed data can provide spatially and temporally varying constraints on the modelling of biophysical landscape variables that are often superior to that achieved by a single static set of model parameters. Data assimilation merges models and observations in a way that compensates for the deficiencies in each (e.g. uncertainty, coverage), resulting in improved accuracy, coverage, and ultimately forecasting capability.

Methods of assimilation are many and varied, however commonalities exist between them. These commonalities are such, that for any time step, the time integrated first guess (the forecast) of soil moisture states are adjusted by an amount determined by the difference between observed and modelled soil moisture (the innovation), which is weighted by the respective error variances of modelled and observed quantities (the gain), to generate revised soil moisture states (the analysis). At this point, the model soil moisture states are out of balance with the other stores and fluxes, until the model integrates forward to the next time step, whereupon water balance is restored through model physics.

In addition to water balance closure, from an operational perspective, is it important that the method of data assimilation be: computationally efficient for routine, automated simulation over the whole model domain; robust to data gaps; and make lasting positive improvements to future predictions of soil water stores and fluxes. Currently, there are very few operational continental land surface modelling systems that provide high-resolution near-real time soil moisture estimates that have been constrained through the assimilation of satellite observations. Some recent examples include surface soil wetness observations from Advanced Scatterometer (ASCAT) active radar system, on the meteorological operational satellite (MetOp), been assimilated into Unified Model (Davies et al., 2005) through nudging to provide soil moisture analysis at 40 km globally (Dharssi et al., 2011). Additionally, ASCAT data are used in the ECMWF (European Centre for Medium-Range Weather Forecasts) Land Data Assimilation System through a simplified Extended Kalman Filter approach (De Rosnay et al., 2013) to provide near-real time surface soil moisture and root-zone soil moisture at 25-km resolution globally. However, soil moisture products from a passive radiometer system such as SMOS (Soil Moisture and Ocean Salinity) mission (Kerr et al., 2001) or the SMAP mission (Entekhabi et al., 2010) have not been fully explored in an operational data assimilation system.

In this study, we develop a simple, computationally efficient, and effective data assimilation approach for assimilating satellite soil moisture products into an operational national water balance model. We demonstrate the application of the method to the Australian Water Resources Assessment Community Modelling system (AWRA-CMS), which provides daily water balance estimates at 5-km resolution across Australia, with the assimilation of satellite surface soil moisture (SSM) from both SMOS and SMAP. A post-analysis adjustment is proposed to impart mass conservation constraints on related states and fluxes such as root-zone soil water storage, evapotranspiration and streamflow thus improving the accuracy of the water balance post assimilation. The impacts of data assimilation on model predictions is assessed by quantifying the persistence of the correction to key model components with respect to open-loop simulations.

2 Materials and Methods

2.1 Australian Water Resources Assessment Community Modelling system (AWRA-CMS)

The Australian Water Resources Assessment (AWRA) Community Modelling system (AWRA-CMS) is a freely available version of the AWRA Landscape model (Van Dijk, 2010) which simulates the water balance in the Australian landscape (https://github.com/awracms/awra_cms). The operational implementation of the AWRA-CMS by the Australian Bureau of Meteorology provides daily 0.05 degree (approximately 5 km) national gridded soil moisture, runoff, evapotranspiration and deep drainage estimates, and underpins the annual national water resource assessments and water use accounts (Frost et al., 2018) as well as providing situational soil moisture for flood forecasting, agriculture and other applications. AWRA is a one-dimensional distributed model that simulates the water balance for each grid cell across the modelling domain by distributing rainfall into plant-accessible water, soil moisture and groundwater stores, and removing water through evapotranspiration, runoff and deep drainage. The soil water column has been partitioned into three layers (upper: 0–10 cm, lower: 10–100 cm, and deep: 1–6 m) simulated separately for deep- and shallow-rooted vegetation. In addition to the modelling of soil columns, the model includes a surface water and a groundwater storage that are simulated at each grid cell and conceptualized as if operating within a small unimpaired catchment. In this study, we used daily precipitation and air temperature from the gridded climate data services (Jones et al., 2009), daily solar exposure produced from geostationary satellites (Grant et al., 2008), and interpolated site-based wind speed (McVicar et al., 2008) as model forcing inputs.

2.2 Satellite soil moisture (SSM)

To optimize the daily spatial coverage, we used two satellite soil moisture products derived from passive L-band systems: the Soil Moisture Active-Passive (SMAP) product from NASA (Entekhabi et al., 2010); and the product from the European Space Agency's (ESA's) Soil Moisture and Ocean Salinity (SMOS) mission (Kerr et al., 2001). The SMAP product is the level-2 enhanced radiometer half-orbit 9-km EASE-grid soil moisture (Chan et al., 2018). The SMOS product is the level-2 soil moisture product on ~ 25-km grid (Rahmoune et al., 2013). Both SMAP and SMOS produce volumetric soil moisture estimates (units m^3/m^3) of approximately the upper 5 cm of soil. Available swath data for each product covering Australia were sourced and collated for 24-hour period approximating the AWRA-CMS operational time steps and interpolated to a regular 0.05-degree grid across the modelling domain from 2015 to 2019. This provided maximum possible spatial coverage for each data product in representing surface soil moisture at the end of the model's time step integration each day.

2.3 Data assimilation approach through triple collocation (TC)

The data assimilation method used here is a time sequential updating of model state(s) given observations of relevant model variables (Reichle, 2008). Two key modelling components in data assimilation: the *dynamics operator*, which describes the time integration of the system states and fluxes, which in this study is the AWRA-CMS; and the *observation operator*, which provides the mathematical mapping from state to observation space (or vice versa). The role of the observation operator is to perform a mapping between observation and state space, as often observations are not directly comparable to model states.

The state updating equation for sequential data assimilation is written as:

$$X_t^a = X_t^f + K_t[Y_t - H(X_t^f)] \tag{1}$$

which says that the best estimate of model state, known as analysis ($\mathrm{X_t^a}$), is equal to the first guess or forecast ($\mathrm{X_t^f}$) plus a weighted difference between observations, $\mathrm{Y_t}$, and the model equivalent to the observation, $\mathrm{H}(\mathrm{X_t^f})$, for that time step. The multiplier, $\mathrm{K_t}$, is known as the *gain factor* which contains uncertainty expressed as error variance for both model estimates (σ_B^2) and observations σ_R^2. For a unity observation operator and assuming independence between model estimates and observations, gain factor typically assumes the form:

$$K = \frac{\sigma_B^2}{\sigma_B^2+\sigma_R^2}\,. \tag{2}$$

In this study, the state variable of focus is the moisture storage in AWRA's upper soil layer, $\mathrm{S_0}$. Satellite surface soil moisture (SSM) products from both SMOS and SMAP are used as the observations to update the model simulation. Satellite soil moisture estimates are provided in volumetric units ($\mathrm{m^3/m^3}$), whereas modeled upper-layer soil moisture is given in terms of storage of water (i.e. units mm). The observation operator used here is a linear transformation which matches the mean and variance between model and observation time series (Tian et al., 2017). As such, the observation operator also simultaneously removes systematic bias between model estimates and satellite observations. In addition, for region with sparse rain-gauge coverage such as central Western Australia, the linear transformation of the satellite soil moisture products draws on data sampled from neighboring cells with similar soil moisture conditions, to account for known poor model estimates from consistent underestimation of rainfall (S1).

The gain factor, K, contains information on the error variances of the model and observations. Observation error variance is often estimated through field campaigns (Draper et al., 2009; Panciera et al., 2014), but these rarely represent the spatial and temporal variability of errors in gridded satellite products. Alternatively, data providers often specify error estimates, but their magnitude can be overly optimistic. Triple collocation (TC) was developed as a method of quantifying error characteristics in geophysical variables when the true error structure is elusive. It was first applied to near-surface wind data (Stoffelen, 1998) and later extensively applied to soil moisture (Dorigo et al., 2017; McColl et al., 2014; Scipal et al., 2008; Su et al., 2014) and rainfall (Massari et al., 2017). The assumption of this approach is that three independent data sets of the same geophysical variable can be used to infer the error variances in each. Here we use TC as a way of inferring error variances from our three independent estimates of surface soil moisture, AWRA $\mathrm{S_0}$, SMAP, and SMOS. McColl et al. (2014) shows that the error variances of each data set can be calculated from the temporal variance and covariance between data sets respectively as:

$$\sigma_x^2 = \left(Q_{x,x} - \frac{Q_{x,y}Q_{x,z}}{Q_{y,z}}\right), \quad \sigma_y^2 = \left(Q_{y,y} - \frac{Q_{x,y}Q_{y,z}}{Q_{x,z}}\right) \quad \text{and} \quad \sigma_z^2 = \left(Q_{z,z} - \frac{Q_{z,y}Q_{x,z}}{Q_{x,y}}\right) \quad (3)$$

where x, y and z denote AWRA, SMAP or SMOS soil moisture estimates respectively and Q denotes temporal variance and covariance between the data sets. These estimates of the error variances are then used in the determination of gain factors (Eq. 2) for the three estimates of soil moisture, thus recasting Equation (1) as:

$$X_t^a = K_{AWRA}X_t^f + K_{SMAP}Y_t^{SMAP} + K_{SMOS}Y_t^{SMOS} \tag{4}$$

The analysed soil water state derived from Eq. (4) represents an optimal blending of AWRA S_0, SMAP and SMOS (S2).

2.4 Analysis increment redistribution (AIR)

The assimilation of satellite soil moisture often violates mass conservation in the model through the analysis update (Eq. 4). The difference between the analysis, X_t^a, and the forecast, X_t^f, (known as the *analysis increment*) represents an amount of water that has been added or subtracted from the system that was not present at the start of model integration for the given time step. In this study, we use the concept of tangent linear modelling (Errico, 1997; Giering, 2000) to redistribute the analysis increment of S_0 to all the relevant model states and fluxes (e.g. lower layer and deep layer soil water storage, evapotranspiration and runoff). This was considered as a way of maintaining mass (i.e. water) balance within a model time step, which data assimilation is known to break. We refer to this approach as analysis increment redistribution, or simply as AIR hereafter. To illustrate, Equation 5 gives an example of what should the resulting changes (Δ) in drainage (D_0) be, given the analysis increment in S_0 ($S_0^a - S_0^f$):

$$\Delta D_0 = (1-\beta_0)k_{0sat}\left[\left(\frac{S_0^a}{S0max}\right)^2 - \left(\frac{S_0^f}{S0max}\right)^2\right], \tag{5}$$

where the $k_{0\mathrm{sat}}$ and S_{0max} are model parameters representing the saturated hydraulic conductivity and maximum storage of the upper soil layer, respectively, and β_0 is the proportion of upper soil layer lateral drainage (S3 for more detail). Corresponding adjustments of total lateral interflow for both upper and lower soil layer are then propagated to the river water storage and total runoff. In addition, the analysis increments of S_0 and change in lower soil layer water storage after application of the AIR are used to revise the total evapotranspiration. The adjustments to the relevant states and fluxes are derived from AWRA model formulation (S3).

2.5 In-situ measurements

Evaluation of the modelled soil water storages was made against measurements from three soil moisture monitoring networks in Australia from 2016 to 2018, namely OzNet (Smith et al., 2012), CosmOz (Hawdon et al., 2014) (Fig. 1a) and OzFlux (Fig. 1b). AWRA model estimates of upper layer soil water storage were compared against in situ measurements from the top 10 cm of soil across all three networks. In situ measurements of root-zone moisture varied across networks from 0-30 cm to 0-1 m. As such, AWRA soil water storages over the root-zone were constructed accordingly by combining upper- and lower-layer soil water storage in the appropriate proportions. OzFlux sites are primarily used for the evaluation of AWRA evapotranspiration estimates, which were calculated from accumulated latent heat flux measurements at each location. In total, there are 45 sites for soil moisture validation and 14 sites for evapotranspiration validation. Streamflow observations for 100 catchments across Australia have been used in the validation based on the quality and data availability (Fig. 1c).

2.6 Vegetation index

In water-limited regions like Australia, shallow-rooted vegetation normally responds quickly with soil water availability, typically within a month. Consistency between root-zone soil water storage and vegetation greenness may be considered as an indirect independent verification of the simulation of root-zone soil water dynamic (Tian et al., 2019a; Tian et al., 2019b). The 0.05-degree monthly Normalized Difference Vegetation Index (NDVI) from Moderate Resolution Imaging Spectroradiometer (MODIS, MYD13C2 v6) is used to evaluate estimates of root-zone soil moisture over cropland and grassland regions of the continent. The 250m land cover classification map from Geoscience Australia (Lymburner et al. 2015) is resampled to 0.05 degree over model domain and used in the identification of crop and grassland cells.

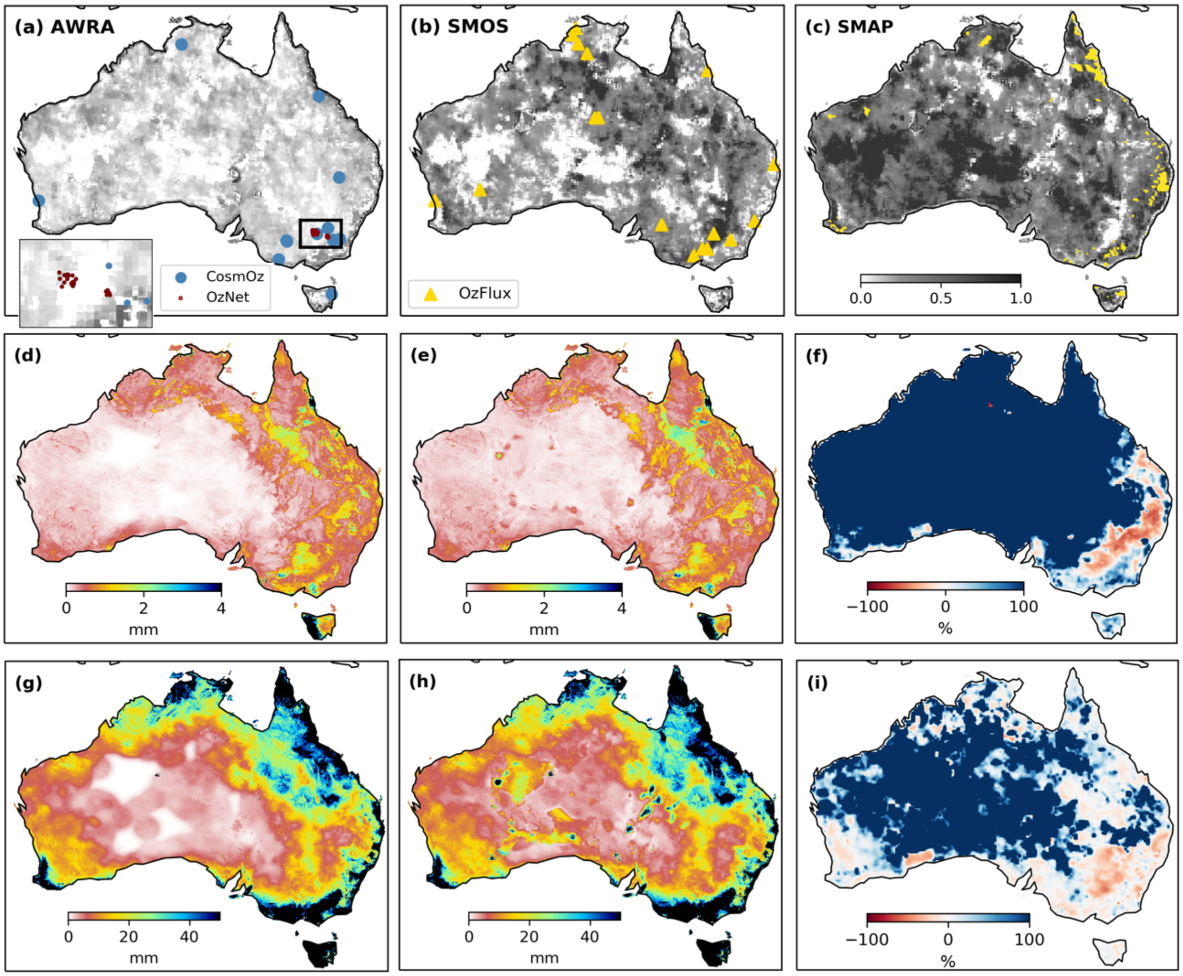

Figure 1 Study area showing gain factors of the TC data assimilation method rescaled to highlight relative contribution of the respective estimate: (a) AWRA-simulated S_0, (b) SMOS soil moisture, and (c) SMAP soil moisture. Also displayed are the locations of in-situ monitoring stations from (a) CosmOz and OzNet networks, (b) OzFlux network, and (c) catchments for streamflow validation. Subfigures (d) and (e) are the average S_0 simulations for 2019 from AWRA open-loop (OL) and TC assimilation of SMOS and SMAP data (DA-TC). Subfigure (f) shows the average relative change of analysed S_0 (TC) compared to OL simulations in 2019. Subfigures (g) and (h) are the average S_s simulations for 2019 from AWRA OL and DA-TC. Subfigure (i) shows the average relative change of analysed S_s after analysis increment redistribution (TC-AIR) compared to OL simulations in 2019.

3 Results and Discussion

3.1 Improved spatial representation of soil moisture variability

The analysed upper layer soil water storage estimates receive a greater contribution from SSM products, in particular SMAP observations, compared to model simulations (Figures 1a-c). AWRA model simulations are driven by gauge-based rainfall analyses. As such they have difficulty in adequately simulating soil moisture patterns over regions lacking in rain gauge coverage, such as Western Australia and central Australia (Fig. 1d). Water storage simulations over these regions default to zero, thus very little or no weight was given to the AWRA estimates in these regions (Fig. 1a). In contrast, SMAP SSM data is heavily weighted in the assimilation due to the smaller error variance derived from TC (Fig. 1c). This is expected since SMAP is the best-performing satellite soil moisture product over the majority of applicable global land pixels (Chen et al., 2018). AWRA simulations of S_0 are dominated by the satellite SSM data as a result of TC data assimilation in the region which largely eliminates the erroneous artefacts associated with deficient rainfall data forcing (Fig.1e).

Moreover, the SSM data assimilation has the effect of adding moisture to AWRA S_0 simulations over most of Australia, with predictions on average often in excess of 100% of those from the OL simulations (Fig. 1f). The notable exception to this is in the southeast of Australia, particularly within the Murray-Darling Basin, where SSM data assimilation reduced AWRA S_0 by more than 50%. This suggests that AWRA simulations underestimated the severity of the drought experienced in the region in 2019.

TC assimilation only updates S_0 directly with satellite SSM, thus the S_s and other water storage receives the impact from assimilation once the model integrates forward to the next time step from the analysed S_0 as initial conditions. The AIR method adjusts S_s and other relevant states and fluxes as a post-correction according to the change in S_0 to maintain water balance. The average S_s with the correction in drainage and lateral interflow from the change in S_0 after TC assimilation shows significant different spatial pattern with a relative change more than 100%

over those regions with sparse rain-gauge coverage against OL simulations (Fig. 1g-i, i.e. see the white regions in Fig1g in particular).

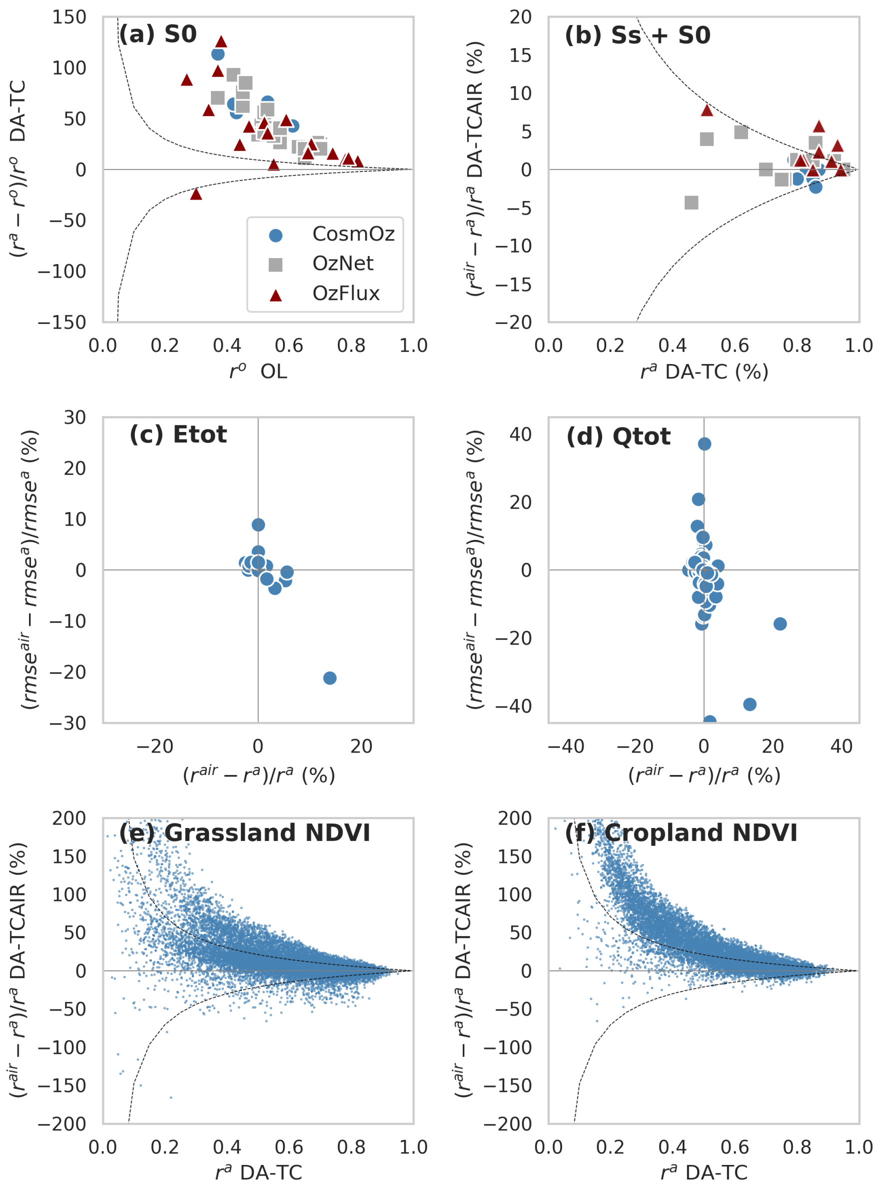

Figure 2 Evaluation of AWRA-estimated upper soil water storage S_0, root-zone soil water storage $S_s + S_0$, total evapotranspiration E_{tot} and runoff Q_{tot}: (a) relative change in correlation of S_0 from TC assimilation (DA-TC) compared to model open-loop (with dots above the zero line showing improved performance); (b) relative change in correlation of root-zone soil water storage from TC-AIR to DA-TC without mass redistribution; (c) – (d) relative correlation and RMSE changes in E_{tot} and Q_{tot} compared to DA-TC without AIR (with dots in the bottom right quadrant showing both improved correlation and reduced RMSE); (e) – (f) relative change in correlation between monthly root-zone soil water storage from TC-AIR with NDVI compared to DA-TC for all grid cells classified as grassland and cropland. Note that dashed curves delineate a 95% level of statistical significance.

3.2 Improved water balance estimates

Comparisons of AWRA simulations with and without SSM data assimilation were made against in-situ measurements networks from 2016 to 2018. Consistent, statistically significant improvement in modelled upper layer soil water storage estimates (S_0) was observed across all sites (Fig. 2a) with the exception of a single OzFlux site located in a tropical rainforest, where microwave SSM retrievals are typically poor in areas of dense vegetation (Njoku and Entekhabi, 1996). TC-based assimilation (Section 2.2) increases the correlation between in-situ surface SM measurements from 0.47 to 0.72 on average for CosmOz sites, 0.54 to 0.69 for OzFlux sites, and 0.56 to 0.77 for OzNet sites compared to OL. This is a significant improvement in AWRA simulations of surface soil moisture dynamics. Compared to ensemble methods of data assimilation (e.g. Tian et al. 2017; 2019b) which rely on an initial guess of the error variance and post hoc correction (e.g. inflation factors, Anderson, 2009), this proposed method based on TC is simple, effective and computationally efficient, thus well suited to an operational system simulating large-scale hydrology. Overall subtle improvements were observed across the AWRA estimates of root-zone soil water storage, evapotranspiration and streamflow (results not shown, see Tian et al., 2019c). The level of improvement is not surprising since those variables were not directly updated with the TC assimilation and are only influenced through the integration of the model to the next time step (Tian et al., 2019c).

The lack of water balance closure is arguably a weak point in data assimilation (Pan and Wood, 2006). Hence, we applied an analysis increments redistribution (AIR, Section 2.3) as a post-correction to all relevant model states and fluxes to enforce mass conservation (water balance). Although the absolute change S_0 is small relative to the volume of S_s, the corresponding change in lower layer soil water storage is allocated through AIR based on model physics (S3). The adjusted root-zone soil water storage (S_0+S_s) shows better agreement with in-situ measurements by up to 10% compared to TC estimates without AIR (Fig. 2b). Improvements are found over the majority of sites from OzFlux and OzNet with measurements. Improvements in correlation together with reduced RMSE (Root-Mean-Squares Error) with in-situ measurements for E_{tot} are found more than 10% relative to the TC estimates without AIR for some sites (Fig. 2c). Further improvements in Q_{tot} simulations are found for some sites, with up to 40% reduction in RMSE (Fig. 2d). Improvements in runoff simulation are due to, first, the SSM assimilation improving pre-storm soil moisture status (Pauwels et al., 2001; Crow and Ryu, 2009), and then AIR adjusts the interflow and river storage accordingly. This indicates the importance of accurate antecedent soil moisture condition in the simulation of runoff response to subsequent rainfall.

The inadequate distribution of in-situ observations as well as the large spatial disparity between ground measurement and modelling scales are a great limitation for the evaluation of root-zone soil moisture and evapotranspiration. AWRA simulation of root-zone soil moisture are compared against satellite-derived NDVI in an indirect verification of model performance and as a way of evaluating the impact of data assimilation. We calculated the correlation between time series of monthly average AWRA root-zone soil moisture from OL, DA-TC and TC-AIR simulations against NDVI for cropland and grassland of Australia over the period 2015 to 2018. These cover types we selected as their rooting depths are commensurate with the combined soil depths of the upper- and lower-soil water storages in AWRA. Figure 2e-f show the relative change in correlation between root-zone simulations from DA-TC and those from TC-AIR data against NDVI data for grassland and cropland areas of Australia. The figure shows that for the vast majority of model grids, TC-AIR shows statistically significant increase in correlation with NDVI compared to DA-TC alone, with an average increase in correlation with NDVI from 0.64 to 0.67 for grassland and 0.55 to 0.66 for cropland compared to OL. This demonstrates that enforcing mass balances as part of the SSM data assimilation each time step is essential to improving the simulation of root-zone soil water balance. The improved consistency with NDVI also illuminates the potential of improving agricultural planning with more accurate information of root-zone soil water availability.

3.3 Impacts on model predictions

Accurate soil water estimates can provide initial conditions for improved flood forecasting and groundwater forecasting (Getirana et al., 2020a; Getirana et al., 2020b; Wanders et al., 2013). Few studies quantify how long the impacts of data assimilation persist in the model system's memory. In this study we used 100-day model simulations from initial states provided by the AWRA OL and DA-TC with AIR. We calculated the number of days it took for the simulation from the analysed DA-TCAIR states to converge to within +/- 5% of those from OL. The experiments were run for one year from 1 March 2018 to 28 February 2019. Results show that data assimilation can impact on model states and fluxes for weeks and sometimes up to 2-3 months (Fig. 3). The impacts of DA-TC with AIR can persist in simulated S_0 for as long as a week over coastal regions, and longer in central Western Australia and Northern Australia with up to 1 month persistence in winter and spring (Fig. 3a). There is less impact on S_0 simulations during wet season since the S_0 can saturate rapidly due to the heavy rainfall. Overall, the longest persistence is found in winter with a continental average of 13 days; the shortest persistence is 6 days on average in autumn and summer. The memory of initial conditions in simulations of S_s can persist even longer due to the slower response to rainfall variability and higher field capacity. Summer persistence for S_s is the least with a continental average of 30 days; in winter, this is increased to 45 days.

Evapotranspiration estimates, however, do not feedback into the system and are highly variable in time and space. On average, the impact of the antecedent soil moisture conditions on evapotranspiration simulations can persist for 1 week over coastal areas, but up to months in central Western Australia. The continental average varies from 13 to 20 days for each season. The areas with the longest persistence are those areas with artefacts of zero rainfall in the forcing. This demonstrates that improvements in AWRA estimates after SSM assimilation over regions with sparse rain-gauge coverage can persist in the system for more than 2 months. The impact on runoff varies from 1 week to 3 months over the continent. The majority of areas impacted for more than 2 months are in locations of little rainfall and runoff. However, there remains between 1-2 week impacts over north-eastern areas with heavy runoff.

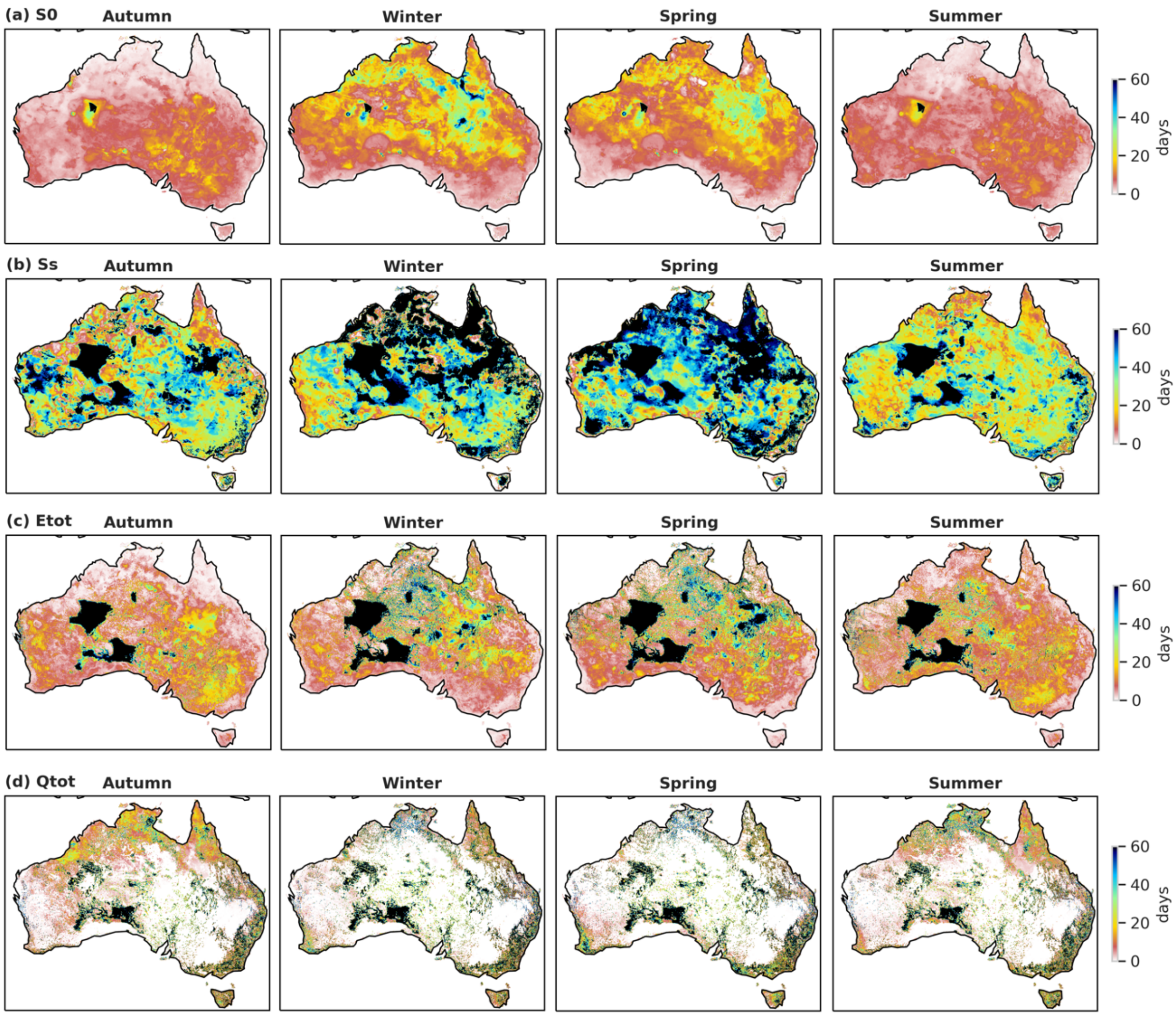

Figure 3 Quantified impacts of data assimilation on forecasting AWRA state variables through the forecast of states for 100 days using the initial condition from DA-TCAIR: average time period that the impact of data assimilation can persist in autumn (2018.03-2018.05), Winter (2018.06-2018.08), Spring (2018.09-2018.11) and Summer (2018.12-2019.02) on (a) upper-layer soil water storage S_0, (b) lower-layer soil water storage S_s, (c) total evapotranspiration E_{tot} and (d) total runoff Q_{tot}.

4 Conclusion

In this study, we proposed a simple and robust method for assimilating SMAP and SMOS soil moisture products into the operational Australian Water Resources Assessment (AWRA) model. The method involves the sequential (daily) updating of the model's upper layer soil water storage with satellite soil moisture observations through a linear combination with weights determined through triple collocation (DA-TC). Evaluation against in-situ measurements showed that simulations of surface soil moisture dynamics is improved significantly after TC data assimilation with an average increase of 0.23 correlation units compared with open-loop simulations. Furthermore, we proposed an additional component to the data assimilation whereby the analysis increment of the upper layer soil water storage is propagated into relevant model states and fluxes as a way of maintaining mass balance (TC-AIR). An evaluation of the root-zone soil moisture, evapotranspiration and streamflow estimates showed that the TC-AIR appeared to only provide marginal, yet positive, improvement over the TC data assimilation method alone. However, in an indirect verification of modelled root-zone soil moisture against satellite-derived NDVI, TC-AIR was seen to provide significant improvement on TC method alone. This demonstrates that by enforcing mass balances as part of the SSM data assimilation each time step, AWRA can better represent soil water dynamics with greater consistency with vegetation response.

The assimilation of satellite soil moisture estimates together with the mass redistribution reduces the uncertainties in model estimates resulting mainly from uncertain forcing and model physics, and provides temporally and spatially varying constraints on model water balance estimates. For example, the assimilation resolves the gaps in rainfall forcing, and the underestimate of drought condition over south-eastern areas in 2019. We demonstrate that the impacts of data assimilation can persist in the model system for more than a week for surface soil water storage and more than a month for root-zone soil water storage. This highlights the importance of accurate initial hydrological states for improving forecast skill over longer lead times. Hence, an operational water balance modelling system, with satellite data assimilation, has strong potential to add value for assessing and predicting water availability for a range of decisions across industries and sectors.

Acknowledgments

This project is supported by collaborative research agreement between the Australian Bureau of Meteorology and Australian National University. We would like to thank Stuart Baron-Hay from the Bureau of Meteorology for his help with implementation of the in AWRA-CMS. This research was undertaken with the assistance of resources and services from the National Computational Infrastructure (NCI), which is supported by the Australian Government through the National Collaborative Research Infrastructure Strategy.

Data Availability

The AWRA-CMS code is accessible from github (https://github.com/awracms/awra_cms). SMAP product used here is the level-2 enhanced radiometer half-orbit 9-km EASE-grid soil moisture from the US National Snow and Ice Data Center (https://nsidc.org). SMOS level-2 soil moisture product is available from ESA's SMOS online dissemination service (https://smos-diss.eo.esa.int/oads/access/). The MYD13C2 NDVI data is accessible through Land Processes Distributed Active Archive Centre (https://lpdaac.usgs.gov). The National Dynamic Land Cover Dataset of Australia is available from Geoscience Australia (https://www.ga.gov.au).

References

Chan, S. K., Bindlish, R., O'Neill, P., Jackson, T., Njoku, E., Dunbar, S., et al. (2018). Development and assessment of the SMAP enhanced passive soil moisture product. *Remote Sensing of Environment, 204*, 931-941.

Chen, F., Crow, W. T., Bindlish, R., Colliander, A., Burgin, M. S., Asanuma, J., & Aida, K. (2018). Global-scale evaluation of SMAP, SMOS and ASCAT soil moisture products using triple collocation. *Remote Sensing of Environment, 214*, 1-13

Crow, W. T., & Ryu, D. (2009). A new data assimilation approach for improving runoff prediction using remotely-sensed soil moisture retrievals. *Hydrology and Earth System Sciences, 13*(1), 1-16.

Davies, T., Cullen, M. J. P., Malcolm, A. J., Mawson, M. H., Staniforth, A., White, A. A., & Wood, N. (2005). A new dynamical core for the Met Office's global and regional modelling of the atmosphere. *Quarterly Journal of the Royal Meteorological Society, 131*(608), 1759-1782.

de Rosnay, P., Drusch, M., Vasiljevic, D., Balsamo, G., Albergel, C., & Isaksen, L. (2013). A simplified Extended Kalman Filter for the global operational soil moisture analysis at ECMWF. *Quarterly Journal of the Royal Meteorological Society, 139*(674), 1199-1213.

Dharssi, I., Bovis, K. J., Macpherson, B., & Jones, C. P. (2011). Operational assimilation of ASCAT surface soil wetness at the Met Office. *Hydrology and Earth System Sciences, 15*(8), 2729-2746.

Dorigo, W., Wagner, W., Albergel, C., Albrecht, F., Balsamo, G., Brocca, L., et al. (2017). ESA CCI Soil Moisture for improved Earth system understanding: State-of-the art and future directions. *Remote Sensing of Environment, 203*, 185-215.

Draper, C. S., Reichle, R. H., De Lannoy, G. J. M., & Liu, Q. (2012). Assimilation of passive and active microwave soil moisture retrievals. *Geophysical Research Letters, 39*.

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

Lymburner, L., Tan, P., McIntyre, A., Thankappan, M., Sixsmith, J. (2015). Dynamic Land Cover Dataset Version 2.1. Geoscience Australia, Canberra. http://pid.geoscience.gov.au/dataset/ga/83868

Entekhabi, D., Njoku, E. G., O'Neill, P. E., Kellogg, K. H., Crow, W. T., Edelstein, W. N., et al. (2010). The Soil Moisture Active Passive (SMAP) Mission. *Proceedings of the Ieee, 98*(5), 704-716.

Errico, R. M. (1997). What is an adjoint model?. *Bulletin of the American Meteorological Society*, 78(11), 2577-2592.

Frost, A., A. Ramchurn, and A. Smith (2018). The Australian Landscape Water Balance model (AWRA-L v6), Technical Description of the Australian Water Resources Assessment Landscape model version 6. *Bureau of Meteorology Technical Report*, Canberra: Bureau of Meteorology.

Getirana, A., Rodell, M., Kumar, S., Beaudoing, H. K., Arsenault, K., Zaitchik, B., et al. (2020). GRACE Improves Seasonal Groundwater Forecast Initialization over the United States. *Journal of Hydrometeorology, 21*(1), 59-71.

Getirana, A., Jung, H. C., Arsenault, K., Shukla, S., Kumar, S., Peters-Lidard, C., et al. (2020). Satellite Gravimetry Improves Seasonal Streamflow Forecast Initialization in Africa. *Water Resources Research, 56*(2), doi: 10.1029/2019wr026259.

Giering, R. (2000). Tangent linear and adjoint biogeochemical models. *Inverse Methods in Global Biogeochemical Cycles, 114*, 33-48.

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, paper presented at Catchment-scale Hydrological Modelling and Data Assimilation (CAHMDA-3). *International Workshop on Hydrological Prediction: Modelling, Observation and Data Assimilation*, Melbourne, Citeseer.

Hawdon, A., McJannet, D., & Wallace, J. (2014). Calibration and correction procedures for cosmic-ray neutron soil moisture probes located across Australia. *Water Resources Research, 50*(6), 5029-5043.

Ines, A. V. M., Das, N. N., Hansen, J. W., & Njoku, E. G. (2013). Assimilation of remotely sensed soil moisture and vegetation with a crop simulation model for maize yield prediction. *Remote Sensing of Environment, 138*, 149-164.

Jones, D. A., Wang, W., & Fawcett, R. (2009). High-quality spatial climate data-sets for Australia. *Australian Meteorological and Oceanographic Journal, 58*(4), 233-248.

Kerr, Y. H., Waldteufel, P., Wigneron, J. P., Martinuzzi, J. M., Font, J., & Berger, M. (2001). Soil moisture retrieval from space: The Soil Moisture and Ocean Salinity (SMOS) mission. *Ieee Transactions on Geoscience and Remote Sensing, 39*(8), 1729-1735.

Koster, R. D., Dirmeyer, P. A., Guo, Z. C., Bonan, G., Chan, E., Cox, P., et al. (2004). Regions of strong coupling between soil moisture and precipitation. *Science, 305*(5687), 1138-1140.

Massari, C., Crow, W., & Brocca, L. (2017). An assessment of the performance of global rainfall estimates without ground-based observations. *Hydrology and Earth System Sciences, 21*(9), 4347-4361.

McColl, K. A., Vogelzang, J., Konings, A. G., Entekhabi, D., Piles, M., & Stoffelen, A. (2014). Extended triple collocation: Estimating errors and correlation coefficients with respect to an unknown target. *Geophysical Research Letters, 41*(17), 6229-6236.

McVicar, T. R., Van Niel, T. G., Li, L. T., Roderick, M. L., Rayner, D. P., Ricciardulli, L., & Donohue, R. J. (2008). Wind speed climatology and trends for Australia, 1975-2006: Capturing the stilling phenomenon and comparison with near-surface reanalysis output. *Geophysical Research Letters, 35*(20)

Njoku, E. G., & Entekhabi, D. (1996). Passive microwave remote sensing of soil moisture. *Journal of Hydrology, 184*(1-2), 101-129.

Pan, M., & Wood, E. F. (2006). Data assimilation for estimating the terrestrial water budget using a constrained ensemble Kalman filter. *Journal of Hydrometeorology, 7*(3), 534-547.

Panciera, R., Walker, J. P., Jackson, T. J., Gray, D. A., Tanase, M. A., Ryu, D., et al. (2014). The Soil Moisture Active Passive Experiments (SMAPEx): Toward Soil Moisture Retrieval From the SMAP Mission. *Ieee Transactions on Geoscience and Remote Sensing, 52*(1), 490-507.

Pauwels, V. R. N., Hoeben, R., Verhoest, N. E. C., & De Troch, F. P. (2001). The importance of the spatial patterns of remotely sensed soil moisture in the improvement of discharge predictions for small-scale basins through data assimilation. *Journal of Hydrology, 251*(1-2), 88-102.

Pipunic, R. C., Walker, J. P., & Western, A. (2008). Assimilation of remotely sensed data for improved latent and sensible heat flux prediction: A comparative synthetic study. *Remote Sensing of Environment, 112*(4), 1295-1305.

Rahmoune, R., Ferrazzoli, P., Kerr, Y. H., & Richaume, P. (2013). SMOS Level 2 Retrieval Algorithm Over Forests: Description and Generation of Global Maps. *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 6*(3), 1430-1439.

Reichle, R. H., & Koster, R. D. (2005). Global assimilation of satellite surface soil moisture retrievals into the NASA Catchment land surface model. *Geophysical Research Letters, 32*(2). https://doi.org/10.1016/j.advwatres.2008.01.001

Reichle, R. H. (2008). Data assimilation methods in the Earth sciences. *Advances in Water Resources, 31*(11), 1411-1418.

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

Scipal, K., Drusch, M. & Wagner, W., (2008). Assimilation of a ERS scatterometer derived soil moisture index in the ECMWF numerical weather prediction system, *Advances in water resource*s, 31(8), 1101-1112.

Sheffield, J., & Wood, E. F. (2007). Characteristics of global and regional drought, 1950-2000: Analysis of soil moisture data from off-line simulation of the terrestrial hydrologic cycle. *Journal of Geophysical Research-Atmospheres, 112*(D17).

Smith, A. B., Walker, J. P., Western, A. W., Young, R. I., Ellett, K. M., Pipunic, R. C., et al. (2012). The Murrumbidgee soil moisture monitoring network data set. *Water Resources Research, 48*.

Stoffelen, A. (1998). Toward the true near-surface wind speed: Error modeling and calibration using triple collocation. *Journal of Geophysical Research-Oceans, 103*(C4), 7755-7766.

Su, C. H., Ryu, D., Crow, W. T., & Western, A. W. (2014). Beyond triple collocation: Applications to soil moisture monitoring. *Journal of Geophysical Research-Atmospheres, 119*(11), 6419-6439.

Tian, S., van Dijk, A. I. J. M., Tregoning, P., & Renzullo, L. J. (2019). Forecasting dryland vegetation condition months in advance through satellite data assimilation. *Nature Communications,* 10. (1), 1-7.

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 Resources Research, 53*(3), 1820-1840.

Tian, S., Renzullo, L. J., van Dijk, A. I. J. M., Tregoning, P., & Walker, J. P. (2019). Global joint assimilation of GRACE and SMOS for improved estimation of root-zone soil moisture and vegetation response. *Hydrology and Earth System Sciences, 23*(2), 1067-1081.

Tian, S., Renzullo, L. J., Pipunic, R., Sharples, W., Lerat, J., & Donnelly, C.(2019),. Assimilating satellite soil moisture retrievals to improve operational water balance modelling. *Proceedings of the 23rd International Congress on Modelling and Simulation*, Canberra, Australia. Https://doi.org/10,36334/modsim.2019.H6.tian

van Dijk, A.I.J.M. (2010). AWRA Technical Report 3, Landscape Model (version 0.5) Technical Description, WIRADA, Canberra: CSIRO Water for a Healthy Country Flagship.

Vinnikov, K. Y., Robock, A., Speranskaya, N. A., & Schlosser, A. (1996). Scales of temporal and spatial variability of midlatitude soil moisture. *Journal of Geophysical Research-Atmospheres, 101*(D3), 7163-7174.

Wanders, N., Karssenberg, D., de Roo, A., de Jong, S. M., & Bierkens, M. F. P. (2014). The suitability of remotely sensed soil moisture for improving operational flood forecasting. *Hydrology and Earth System Sciences, 18*(6), 2343-2357.