

Enhancement of water storage estimates using GRACE data assimilation with particle filter framework

N. Tangdamrongsub, S.-C. Han and I.Y. Yeo

School of Engineering, University of Newcastle, Callaghan, NSW 2308, Australia

Email: Nathachet.tangdamrongsub@newcastle.edu.au

Abstract: An accurate knowledge of soil moisture and groundwater storage is crucial to understand hydrological process and extreme climate events. The model outputs of the terrestrial water storage are biased by inaccurate forcing data, inefficacious model physics, and improper model parameter calibration. To mitigate the model uncertainty, the observation (e.g., from remote sensing as well as ground in-situ data) are often integrated into the model to improve the simulation result via data assimilation (DA).

This study intends to enhance the estimation of soil moisture and groundwater storage by assimilating the Gravity Recovery And Climate Experiment (GRACE) observation into the Community Atmosphere Biosphere Land Exchange (CABLE) land surface model using the particle filter (PF) framework. The PF is developed for GRACE DA in order to accommodate different types of posterior error distribution and thus allow the realistic system representation where the distribution of model and observation errors are usually unknown. The early development of PF commonly suffered from the particle degeneracy and impoverishment problems, mainly caused by the insufficient number of particles. This study uses the sequential importance resampling (SIR) approach to reduce the problems. The simulation conducted to evaluate the filter performance and determine the effective number of particles shows that the SIR approach can deliver the accurate water storage estimates with the usage of only 100 particles. Moreover, the uncertainty of GRACE observation is obtained directly from the full error variance-covariance matrix provided as a part of the GRACE data product. This method demonstrates the use of a realistic representative of GRACE uncertainty, which is spatially correlated in nature, leads to an improvement of storage computation.

The developed GRACE DA scheme is demonstrated over the Goulburn catchment located in the Upper Hunter region, NSW, where the ground observations (surface soil moisture, root-zone soil moisture, and groundwater level) from the Scaling and Assimilation of Soil Moisture and Streamflow (SASMAS) network and the Department of Primary Industries, Office of Water, New South Wales are available for evaluation of our DA results. This study is the first time the GRACE-PF is exploited in a small catchment size ($\leq 6,540 \text{ km}^2$), proving an important insight about the potential of GRACE over a smaller region beyond its limit of spatial resolution at $\sim 250 \text{ km}$.

Preliminary results show that our developed technique successfully disaggregates the catchment-scale GRACE information into finer vertical and spatial scale ($\sim 25 \text{ km}$), leading to a significant improvement particularly in groundwater and, marginally in deep soil moisture components. On average, GRACE DA improves the groundwater storage computation in terms of correlation coefficient (ρ) by approximately $\sim 47\%$ (from 0.38 to 0.56). The ρ value changes from 0.535 to 0.543 by only 1.4% for the deeper soil moisture (beneath 60 cm) computation. The improvement is found mainly from deeper layers with slower temporal variations, which is consistent with the interannual time scale of the GRACE signals being most characteristic over that catchment. However, GRACE DA slightly degrades the computation of the near surface soil moisture by approximately 2.2% (in ρ). The coarse temporal and spatial resolution of GRACE is attributed to the less impact of the GRACE DA on surface soil moisture estimation.

In conclusion, it is apparent that GRACE DA provides a crucial benefit to deep storage computation. Further development will incorporate satellite soil moisture observations from Soil Moisture Ocean Salinity (SMOS) and Soil Moisture Active Passive (SMAP) missions with GRACE in the assimilation scheme to simultaneously improve different storage components, including surface soil moisture. Comprehensive evaluation of PF's results in comparison to EnKF results will also be conducted to understand the filter's performance with regard to accuracy of water storage estimates.

Keywords: GRACE, data assimilation, particle filter, CABLE, Goulburn catchment, SASMAS, groundwater, soil moisture

1. INTRODUCTION

The terrestrial water storage (TWS) is a vital information for comprehensive understanding of the climate variation and the accurate assessment of the regional/global water resources. The TWS can be defined (but not limited to) as the integration of soil moisture, groundwater, surface water, snow, and canopy water storage, which all undoubtedly play very important roles in the hydrological cycle. The TWS can be simulated from the land surface model but the accuracy of computation is generally governed by the quality of the model physics and parameter calibration, and the accuracy of the input meteorological data. An alternative approach is to derive the TWS from Gravity Recovery And Climate Experiment (GRACE) observations, which has shown to be effective at a regional scale (Famiglietti *et al.*, 2011). This motivates the development of GRACE data assimilation (DA) to improve the model estimates by integrating the GRACE observation into the system.

GRACE DA has been effectively exploited to improve various hydrological components, e.g., groundwater (Tangdamrongsub *et al.*, 2015), snow water equivalent (Su *et al.*, 2010), root zone soil moisture and latent heat flux (Zaitchik *et al.*, 2008), etc. Various DA schemes have recently been developed, including the inclusion of the GRACE's full error variance-covariance information (e.g., Tangdamrongsub *et al.*, 2017). However, most GRACE DA frameworks were developed based the ensemble Kalman filter (EnKF), which assumes the posterior distribution of the estimates to be Gaussian. In the extreme non-linear system, such as hydrologic system, however, the assumption might be violated and the DA might lead to the incorrect state estimate (Moradkhani *et al.*, 2005). Particle filter (PF) is a variant of DA that does not restrict the type of the posterior distribution, which in turn can lead to a more reliable representation of the system. The PF has been successfully applied in the soil science (Dong *et al.*, 2015), streamflow estimation (Weerts and El Serafy 2006), and only recently exploited in the GRACE literature (Khaki *et al.*, 2017).

In this study, GRACE data is assimilated into the Community Atmosphere Land Exchange (CABLE, Decker 2015) model to improve the soil moisture and groundwater estimates. Our developed GRACE DA is demonstrated over the Goulburn river catchment located in the Upper Hunter Region of NSW, where the dense ground observation data are available for the validation. Importantly, this study is the first time the GRACE-PF is exploited in such a small catchment ($\leq 6,540 \text{ km}^2$), proving an insight about the potential of GRACE over a region size beyond its limited spatial resolution of $\sim 250 \text{ km}$. The performance of the developed PF scheme is evaluated against the ground observation data (volumetric soil moisture and groundwater level variation).

2. STUDY REGION

The Goulburn river catchment is located at the south-eastern part of the Murray-Darling basin and has temperate and semiarid climate (Fig. 1). The catchment is maintained by the Scaling and Assimilation of Soil Moisture and Streamflow (SASMAS) project (<http://www.eng.newcastle.edu.au/sasmas/SASMAS/sasdata.html>), and has been used for soil science, meteorology, and streamflow prediction. The Goulburn catchment has a total area of $6,540 \text{ km}^2$ and consists of more than 10 sub-catchments, including Krui and Merriwa catchments where the in situ soil moistures are regularly recorded.

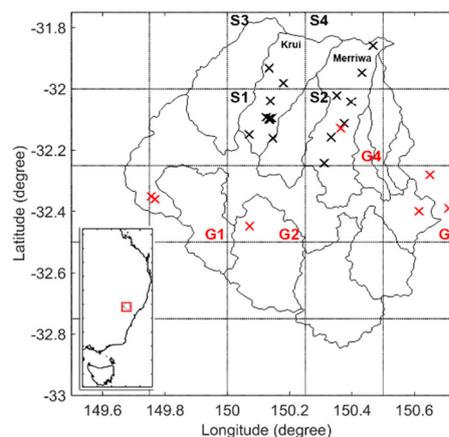


Figure 1. The geolocation of the Goulburn catchment (see also the inset figure (bottom-left)). The model grid cell ($\sim 25 \text{ km}$) is shown as a square tiles. The black x (cross) represents the location of the in situ soil moisture measurement while the red x (cross) represents the groundwater bore. All in situ data inside the same model grid cell are averaged. The grids S1 – S4 represent the averaged in situ soil moisture grid cells while G1 – G4 represent the averaged in situ groundwater cells.

3. DATA AND LAND SURFACE MODEL

3.1. GRACE Observation

GRACE data and its full error variance-covariance matrix between January 2003 and December 2015 are obtained from the Center for Space Research (CSR), the University of Texas Austin. The error matrices beyond June 2014 are not available, and the monthly average values are used for the missing months. GRACE-derived TWS variation (Δ TWS) and its uncertainty over the Goulburn catchment is computed following the approach in Tangdamrongsub *et al.* (2017). In brief, the destriping and smoothing filters are applied to the GRACE's gravity coefficients before computing the Δ TWS. The signal restoration is applied to compensate the signal loss by the filters. To be consistent with the model estimate, the temporal mean value of TWS (Jan 2003 – Dec 2015) from the model estimates is added to GRACE Δ TWS to obtain the absolute TWS. Additionally, the GRACE Δ TWS error matrix is derived based on its full error-variance covariance matrix using the error propagation law. This approach represents a more realistic GRACE uncertainty, which is spatially correlated, compared to the approach of the off-diagonal terms being omitted.

3.2. In-situ Soil Moisture and Groundwater Data

The surface and root zone soil moisture are obtained from the SASMAS network. Four different data associated with the measured depth at 0 – 5, 0 – 30, 30 – 60, and 60 – 90 cm are available in terms of volumetric soil moisture (θ). Furthermore, the in situ groundwater levels are obtained from the Department of Primary Industries, Office of Water, NSW (<http://www.water.nsw.gov.au>). All data inside the same model grid cell (Fig 1) are then averaged for the later comparison. This results in four grid data of in situ soil moisture (S1 – S4) and four of the groundwater (G1 – G4). In situ data are averaged from hourly/daily to monthly time span for the later comparison with the GRACE DA result.

3.3. CABLE Land Surface Model

The Community Atmosphere and Biosphere Land Exchange (CABLE) model is used to predict the volumetric soil moisture and groundwater storage at approximately 25 km resolution. The model grid distribution is shown in Fig. 1. In this study, the model is forced with the meteorological input from the Global Land Data Assimilation System but replaced the precipitation with the data from Tropical Rainfall Measuring Mission. In DA process, precipitation, shortwave radiation, and air temperature are perturbed using an additive white noise with 10% of the nominal value. Model parameters associated with soil moisture and groundwater components are also perturbed with the same size (10%).

4. GRACE DATA ASSIMILATION

4.1. Particle Filter

Particle filter uses the Monte Carlo method and associated weights (of each particle) to approximate the probability density function of the states' estimate. Various implementation of PF has been developed to reduce the particle degeneracy and impoverishment problems (Weerts and El Serafy, 2006). This study uses the sequential importance resampling (SIR) approach. The GRACE DA process is shown in Fig. 2. The model states of each particle are firstly propagated through the model to obtain the forecasted or predicted states. When the observation is available, the states are updated using the new weights computed based on the likelihood function (see the right part of Fig. 2). The weight of each particle (w_t^i) at time t is computed as follows:

$$w_t^i = w_t^{i*} / \sum_{i=1}^N w_t^{i*}, \quad (1)$$

$$w_t^{i*} = w_{t-1}^{i*} \frac{1}{(2\pi)^{M/2} \det(\mathbf{R})^{1/2}} e^{\left[-0.5(\mathbf{y}_t - \mathbf{H}\mathbf{x}_t^i)^T \mathbf{R}^{-1}(\mathbf{y}_t - \mathbf{H}\mathbf{x}_t^i)\right]}, \quad (2)$$

where \mathbf{y} is the observation vector containing GRACE observation, \mathbf{x} is the model state vector, \mathbf{H} is the operational matrix transforming the model state vector to TWS, \mathbf{R} is the error matrix of the observation, M is the number of observation, and i is the index of the particle. After a few updates, most weights might degenerate to negligible values (e.g., close to zeros), and the propagations of these particles are meaningless. To resolve the severe depletion of particle, the resample approach is used to resample the model particles (Weerts and El

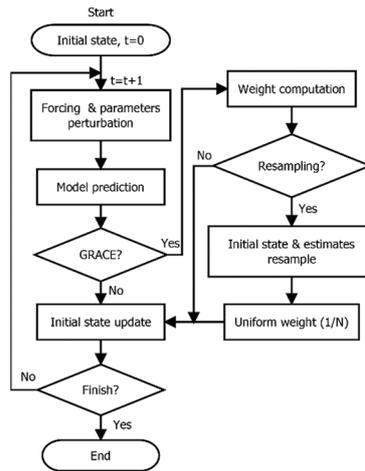


Figure 2. The processing diagram of the GRACE-PF scheme used in this study.

Serafy, 2006). The resampled particles are then propagated to the next time step and the DA process is repeated until the last time step of the study period.

4.2. Implementation

The model states are updated at monthly time scale in consistence with GRACE temporal resolution. The model state vector (\mathbf{x}) contains all grid cells of daily volumetric soil moisture in 6 different layers and groundwater components (total 7 elements) within approximately 1 month. The length of the vector is $J \times T \times 7$, where J is the number of grid cells in the study area, and T is the number of day in a month. The monthly time window used for each update is based on the time tag of GRACE product. As the monthly window used to produce GRACE solution is not necessarily a calendar month (e.g., 1 Jan – 31 Jan), T is different in each update and varies between 13 and 31 days (following GRACE data used). Additionally, GRACE-derived ΔTWS is almost spatially uniform over the study area. The observation vector \mathbf{y} then contains the monthly average values of the catchment mean TWS. The matrix \mathbf{H} contains the scalar operator used to convert the volumetric soil moisture and groundwater into TWS of the month. The elements of \mathbf{H} are simply the thickness of each soil layer (2.2, 5.8, 15.4, 40.9, 108.5, 287.2 cm) and groundwater (20 m). For clarity, the matrix \mathbf{H} is demonstrated below:

$$\mathbf{H} = [\mathbf{h}_{d=1} \quad \mathbf{h}_{d=2} \quad \cdots \quad \mathbf{h}_{d=T}] \quad (3)$$

$$\mathbf{h}_d = [\mathbf{g}_{j=1} \quad \mathbf{g}_{j=2} \quad \cdots \quad \mathbf{g}_{j=J}] \quad (4)$$

$$\mathbf{g}_j = [s_1 \quad s_2 \quad s_3 \quad s_4 \quad s_5 \quad s_6 \quad gw] / JT \quad (5)$$

where d represents the index of day, j is the index of grid cell, s is the thickness of each soil layer, and gw is the thickness of the groundwater layer.

5. RESULTS

5.1. Effective Number of Particles

It is known that the number of particles is an important factor of the filter performance. We investigate this by evaluating the performance of the filter based on different number of particles. We firstly simulate the TWS from CABLE based on different forcing data and parameter sets, and define it as the truth. The error computed based on full error variance-covariance matrix of GRACE is used as the observation error. We then perform the DA (Fig. 2) using 10, 20, 50, 100, 200 and 300 particles. The catchment-average TWS from DA is then compared with the truth in terms of Nash-Sutcliffe (NS) coefficient, and the results are shown in Fig. 3. The figure shows that applying PF leads to a significant agreement with the truth compared to the ensemble open loop (or model-only, EnOL) result. The filter performance is improved when a larger number of particles is used. Figure 3 also shows that the application of approximately 100 particles are sufficient in this study as no further improvement is seen when a greater number of particles is used. However, as the computation effort is not an issue in our study, we perform the GRACE DA using 300 particles to ensure the effectiveness of the filter.

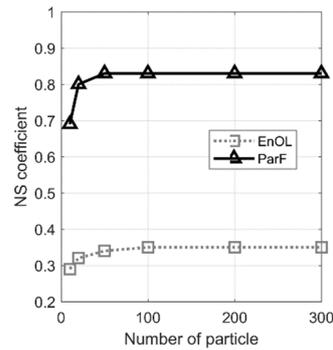


Figure 3. NS coefficient between the catchment-averaged model-simulated TWS and the truth when different number of particle is used. EnOL represents the model-run only while PF is the particle filter result.

5.2. TWS Estimates from GRACE DA

The TWS estimates before and after GRACE DA are evaluated. The catchment-averaged Δ TWS is used for the comparison (Fig. 4). The PF result agrees better with the GRACE observation than the EnOL result. It indicates that the PF approach successfully assimilates GRACE information into the final Δ TWS estimate. The NS coefficients estimated between 2003 and 2015 is substantially improved from 0.3 (EnOL) to 0.7 (PF). In Fig. 4, GRACE observations show greater variation at the beginning period (e.g., 2003 – 2004) which is likely caused by greater observation error at the beginning phase of the GRACE mission. As such, the PF estimated TWS is relatively less influenced by the GRACE observation during this period (NS values from PF in 2003-2004 is only 0.3). The Δ TWS noticeably moves toward GRACE for the remaining period, particularly in e.g., 2006/2007, 2008, and 2014. The model simulates lowered TWS in 2006/2007 by approximately 54 %, and the estimate is corrected after assimilating GRACE. Similar behaviour is seen when GRACE DA removes approximately 3 cm and 6 cm of water storage from the catchment when the model overestimates TWS in 2008 and 2014, respectively. It is worth noting that GRACE DA changes the model storages in several extreme climate events. These include the increasing TWS during 2006/2007 drought, adding TWS during 2012 La Niña period, and subtracting TWS after the La Niña phase (2013 – 2014).

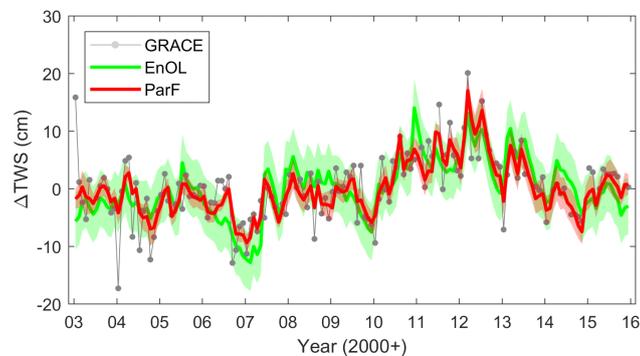


Figure 4. Catchment-averaged Δ TWS estimated from the GRACE observation, model-only (EnOL), and GRACE DA (PF). The standard deviation is shown as the shading background.

5.3. Improvement of Groundwater Storage Estimates

The groundwater storage estimates are validated against the in situ groundwater records at four grid locations (G1 – G4, see Fig. 1). As the specific yield values are not available, we do not convert the in situ data (GW level change, Δ H) to the storage unit, and only validate the results in terms of temporal correlation. Time series and correlation coefficients (ρ) of the EnOL and PF groundwater estimates at G1 – G4 are shown in Fig. 5. Similar to Δ TWS, the PF estimated Δ GWs moves toward the in situ data. Particularly, GRACE DA improves the correlation in all locations, and on average, it increases the correlation value by \sim 47 % (from 0.38 to 0.56). Significant improvements are seen at G2 and G4, particularly at G4 where the correlation value increases almost twice (from 0.35 to 0.68). The good agreement is due to the slow hydrological process of the deep storage. The water passes through soil layers, acting like a low-pass filter, before reaching the groundwater layer. This leads to slower temporal variations of groundwater (compared to the top layers), which is consistent

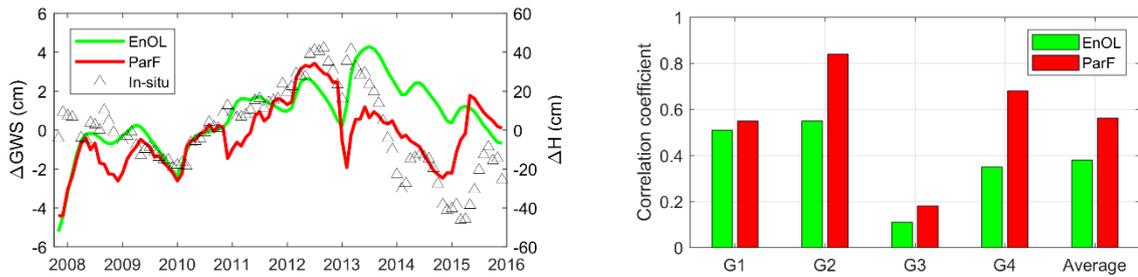


Figure 5. Left: catchment-averaged Δ GWS and in situ data (GW level change, Δ H). Right: correlation coefficient between the groundwater storage estimates (from EnOL, and PF) and the in situ data in four different grid locations (see Fig. 1). The averaged correlation values are also shown.

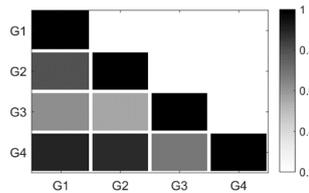


Figure 6. Cross correlation computed from the in situ groundwater data from four different grid locations (G1 – G4). The correlation value is dimensionless.

with the interannual time scale of the GRACE signal. Additionally, GRACE DA apparently benefits the groundwater estimates of the Goulburn catchment, particularly at a finer spatial resolution (\sim 25 km) compared to GRACE itself (\sim 250 km). This is possible likely due to the large unconfined aquifer of the Goulburn catchment. The groundwater variation in finer grid cells are spatially correlated (Fig. 6). The spatial correlation value is always greater than 0.55 and can reach as high as 0.9 (e.g., G1 vs. G4). In such a condition, assimilating a coarser spatial scale Δ TWS from the GRACE observation can benefit the groundwater estimate in the smaller individual grid cells.

5.4. Performance of Soil Moisture Estimates

The soil moisture estimates from GRACE DA is validated against the in situ data. The measured depth between the in situ data and the estimate is different, and therefore we aggregate the soil moisture estimates to have the measured depth as close as the in situ data. For example, the first two layers of CABLE estimated soil moisture (0 – 8 cm) are aggregated before comparing with the in situ surface soil moisture θ_1 (0 – 5 cm). However, the bias due to the different measured depth might still exist, and therefore we only use the correlation coefficient to quantify the goodness of fit here. The correlation coefficient of $\theta_1 - \theta_4$ is firstly computed at each grid location (S1 – S4), and the averaged correlation values (from all S1 – S4) are shown in Fig. 7. Noticeably, no substantial change is found from GRACE DA (2 % degradation from the near surface soil moisture estimates $\theta_1, \theta_2, \theta_3$ and 1% improvement from θ_4 , as opposed to EnOL). This is likely due to the fact that the near surface soil moisture exhibits more immediate responses to precipitation and evapotranspiration than what GRACE monthly measurement can capture. In addition, there is a greater spatial variability of the near surface soil moisture. Such a difference in temporal and spatial characteristics makes GRACE DA less influence surface storage computation.

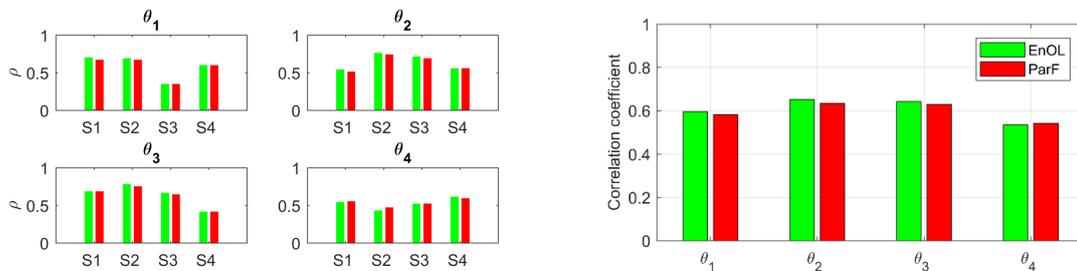


Figure 7. Left: correlation coefficients of four different soil moisture layers ($\theta_1 - \theta_4$) computed between the groundwater storage estimates (from EnOL, and PF) and the in situ data at S1 – S4 (see Fig. 1). Right: the averaged correlation coefficients computed using correlation values from four different locations.

6. CONCLUSIONS

This study assimilates GRACE-derived TWS observation into the CABLE model using the PF framework, and the implementation is demonstrated over the Goulburn catchment. The GRACE DA helps to disaggregate the catchment-scale GRACE information into much finer vertical and spatial scale. The GRACE DA significantly improves computation of groundwater and deep soil moisture (beneath 60 cm) while it does not essentially change surface soil moisture computation. The different temporal and spatial characteristics of GRACE data make DA less responsive to surface soil moisture. Apparently, GRACE DA provides substantial benefit to the deep storage component.

To further evaluate the advantage of PF, the GRACE-PF estimates will be evaluated against the results from other techniques, e.g., GRACE-EnKF. This will provide more insight onto different DA approaches, and also help maximize the accuracy of state estimation. Furthermore, the combined DA of satellite soil moisture observations from Soil Moisture Ocean Salinity (SMOS), and Soil Moisture Active Passive (SMAP) missions with GRACE will be developed with the aim of improving all storage components, including the surface soil moisture (Tian *et al.*, 2017). Finally, the developed DA will be applied to various river catchments to evaluate the effectiveness of the technique under the different climate condition and land cover.

ACKNOWLEDGMENTS

This work is funded by University of Newcastle to support NASA's GRACE and GRACE Follow-On missions. GRACE data are obtained from <http://www.csr.utexas.edu/grace>. The CABLE model can be obtained after the registration at <https://trac.nci.org.au/trac/cable/wiki>.

REFERENCES

- Decker, M. (2015). Development and evaluation of a new soil moisture and runoff parameterization for the CABLE LSM including subgrid-scale processes, *Journal of Advances in Modeling Earth Systems*, 7, 1788–1809, doi:10.1002/2015MS000507.
- Dong, J., Steele-Dunne, S.C., Judge, J., and van de Giesen, N. (2015). A particle batch smoother for soil moisture estimation using soil temperature observations, *Advances in Water Resources*, 83, 111-122.
- Famiglietti, J.S., Lo, M., Ho, S.L., Bethune, J., Anderson, K.J., Syed, T.H., Swenson, S.C., de Linage, C.R., and Rodell, M. (2011). Satellites measure recent rates of groundwater depletion in California's Central Valley, *Geophysical Research Letter*, 38, L03403, doi:10.1029/2010GL046442.
- Khaki, M., Hoteit, I., Kuhn, M., Awange, J., Forootan, E., van Dijk, A.I.J.M., Schumacher, M., and Pattiaratchi, C. (2017) Assessing sequential data assimilation techniques for integrating GRACE data into a hydrological model, *Advances in Water Resources*, 107, 301 - 316.
- Su, H., Yang, Z.-L., Dickinson, R.E., Wilson, C.R., and Niu, G.-Y. (2010). Multisensor snow data assimilation at the continental scale: The value of Gravity Recovery and Climate Experiment terrestrial water storage information, *Journal of Geophysical Research*, 115, D10104, doi:10.1029/2009JD013035.
- Tangdamrongsub, N., Steele-Dunne, S.C., Gunter, B.C., Ditmar, P.G., and Weerts, A.H. (2015). Data assimilation of GRACE terrestrial water storage estimates into a regional hydrological model of the Rhine River basin, *Hydrology and Earth System Sciences*, 19, 2079-2100, doi:10.5194/hess-19-2079-2015.
- Tangdamrongsub, N., Steele-Dunne, S.C., Gunter, B.C., Ditmar, P.G., Sutanudjaja, E.H., Xie, T., and Wang, Z. (2017) Improving estimates of water resources in a semi-arid region by assimilating GRACE data into the PCR-GLOBWB hydrological model, *Hydrology and Earth System Sciences*, 21, 2053 – 2074, doi:10.5194/hess-21-2053-2017.
- Tian S., Tregoning P., Renzullo L.J., van Dijk A.I.J.M., Walker J.P., Pauwels V.R.N., and Allgeyer S. (2017). Improved water balance component estimates through joint assimilation of GRACE water storage and SMOS soil moisture retrievals, *Water Resources Research*, 53, doi:10.1002/2016WR019641.
- Weerts, A.H., and El Serafy, G.Y.H. (2006). Particle filtering and ensemble Kalman filtering for state updating with hydrological conceptual rainfall-runoff models, *Water Resources Research*, 42, W09403, doi:10.1029/2005WR004093.
- Zaitchik, B.F., Rodell, M., and Reichle, R.H. (2008). Assimilation of GRACE terrestrial water storage data into a land surface model: Results for the Mississippi River Basin, *Journal of Hydrometeorology*, 9, 535–548, doi:10.1175/2007JHM951.1.