Water Resources Research

RESEARCH ARTICLE
10.1029/2018WR022785

Key Points:
- It is important to carefully select the EnKF structure, when assimilating GRACE TWS
- TWS change assimilation can be an alternative to TWS when an unbiased TWS baseline is not available
- The impact of assimilation is most pronounced on the largest and most uncertain storages

Supporting Information:
- Supporting Information S1

Correspondence to:
A. Shokri,
ashkan.shokri@monash.edu

Citation:
Shokri, A., Walker, J. P., van Dijk, A. I. J. M., & Pauwels, V. R. N., (2018). Performance of different ensemble Kalman filter structures to assimilate GRACE terrestrial water storage estimates into a high-Resolution hydrological model: A synthetic study. *Water Resources Research*, *54*, 8931–8951. https://doi.org/10.1029/2018WR022785

Received 15 FEB 2018
Accepted 9 OCT 2018
Accepted article online 19 OCT 2018
Published online 14 NOV 2018

Performance of Different Ensemble Kalman Filter Structures to Assimilate GRACE Terrestrial Water Storage Estimates Into a High-Resolution Hydrological Model: A Synthetic Study

Ashkan Shokri[1], **Jeffrey P. Walker**[1], **Albert I. J. M. van Dijk**[2], **and Valentijn R. N. Pauwels**[1]

[1]Department of Civil Engineering, Monash University, Clayton, Victoria, Australia, [2]Fenner School of Environment and Society, Australian National University, Canberra, ACT, Australia

Abstract Among all remote sensing missions, the Gravity Recovery and Climate Experiment (GRACE) was unique as it measured the change in total water content across all terrestrial water storages (TWS) including subsurface, deep soil moisture, and groundwater. However, its coarse resolution is a major challenge for practical applications. Ensemble Kalman filters (EnKFs) are useful tools to combine observations with models to reduce prediction errors. But due to the coarse resolution of the GRACE products, the EnKF does not work well in its usual form. Accordingly, different EnKF structures have been proposed and employed but a comparison between them has not yet been attempted. Here we assessed these structures using a synthetic problem. Alternative structures were formed using different increment calculation and updating strategies, observation operators, and the types of observation fed to the filter. It was found that all available structures led to an improvement in model performance when measured against a synthetic reference. However, the degree of improvement was strongly dependent on the assimilation strategy. Assimilating absolute TWS values (the summation of the TWS anomalies and an unbiased baseline) gave the best model performance when combined with an increment calculation strategy in which the increments are calculated and applied to all days of the month. However, without an unbiased baseline, assimilating TWS changes still leads to an acceptable improvement in model performance. Among the observation operators, those that predict the observations as an average of multiple days had the best performance.

1. Introduction

The world's available freshwater supply is rapidly decreasing, with around 80% of the world's population experiencing high levels of water security threats (Vörösmarty et al., 2010). Consequently, detailed information on water availability is crucial for effective water management. Different water storages including subsurface, deep soil moisture, and groundwater make up the important water resources. Knowledge about these storages is critical to providing useful insights into atmospheric processes, climate change, agriculture productivity, and flooding (Rodell et al., 2004; Tapley et al., 2004). For instance, groundwater is a crucial water supply in some regions and soil moisture has considerable effects on the performance of general circulation models (Entekhabi et al., 1996; Robock et al., 1998; Tapley et al., 2004). Therefore, accurate estimates of the water content of these storages can provide a better understanding about the current state of the hydrologic cycle, improve its future predictability, and consequently help water managers, farmers, and many other stakeholders. However, due to the high heterogeneity, water resources estimation can be challenging (De Lannoy & Reichle, 2016; Hirschi et al., 2014).

To estimate the water content of different water stores, two main approaches exist: observation and hydrological simulation (Tapley et al., 2004). The advance of remote sensing technology has increased our ability to observe water availability at global scale. Among all remotely sensed observations the Gravity Recovery and Climate Experiment (GRACE) mission is the only satellite to provide overall information on the total water content in the form of terrestrial water storage (TWS) change over time. Other water-related missions just observe moisture in the top few centimeters of soil, the water level in lakes, or the snow water equivalent. Therefore, the GRACE mission provides a unique opportunity to explore the Earth's available water with more insight (Rodell et al., 2009; Strassberg et al., 2009; Syed et al., 2009; Tang et al., 2010; Wang et al., 2011; Yeh et al., 2006). However, the coarse temporal (monthly) and spatial (~150,000 km^2) resolution, together with

the vertical aggregation (Rowlands et al., 2005; Swenson & Wahr, 2006), complicate the use of the GRACE TWS retrievals (Girotto et al., 2016).

The hydrological modeling approach for estimation of TWS suffers from inevitable structural and random uncertainties. Errors in the model prediction are propagated at each time step, which can severely affect the accuracy of the final TWS estimates (Ellett et al., 2006).

The ensemble Kalman filter (EnKF) is a very useful data assimilation (DA) tool to reduce the model errors at each time step by updating the model state variables (Evensen, 2003). DA methods have been successfully applied in hydrologic modeling to assimilate observations from various sources into the model, including soil moisture (e.g., Aubert et al., 2003; Crow & Ryu, 2009; Houser et al., 1998; Pauwels et al., 2001; Reichle et al., 2008, 2004), snow water equivalent (e.g., Barrett, 2003; Slater & Clark, 2006; Sun et al., 2004), streamflow (e.g., Clark et al., 2008; Lee et al., 2011), groundwater levels (e.g., Hendricks Franssen et al., 2017), microwave radiances (e.g., Dechant & Moradkhani, 2011), and TWS (e.g., van Dijk et al., 2014; Ellett et al., 2006; Forman & Reichle, 2013; Forman et al., 2012; Girotto et al., 2016, 2017; Houborg et al., 2012; Khaki, Ait-El-Fquih, et al., 2017; Khaki, Hoteit, et al., 2017; Khaki, Schumacher, et al., 2017; Kumar et al., 2016; Li & Rodell, 2015; Li et al., 2012; Smith, 2013; Tian et al., 2017; Zaitchik et al., 2008). Therefore, assimilating the GRACE TWS retrievals into a hydrological model yields more reliable water storage estimates in which the drawbacks of both approaches are mitigated.

To achieve the most accurate results possible, an optimized application of the EnKF is needed. However, due to the temporally and spatially coarse resolution and vertical aggregation of the GRACE TWS retrievals, the regular EnKF structure cannot be used. A number of previous studies have proposed alternative customized EnKF structures for the GRACE TWS retrievals assimilation. The earliest approach was a *two-step* assimilation scheme employed by Zaitchik et al. (2008) to assimilate the GRACE TWS retrievals into the Catchment Land Surface Model. At each time step, the model was integrated to calculate the forecast state variables, observation predictions, and monthly averaged analysis increments. The increments were then broken down to the number of days and gradually added to the state variables to update the model prognoses. This approach has also been used in other studies (Forman & Reichle, 2013; Forman et al., 2012; Houborg et al., 2012; Kumar et al., 2016; Li & Rodell, 2015; Li et al., 2012). Alternatively, different studies applied a regular EnKF in which just the state variables of the last day of the month were updated (Eicker et al., 2014; Schumacher et al., 2016; Su et al., 2010; Tangdamrongsub et al., 2015).

Girotto et al., (2016, 2017) proposed another DA structure. The increments for each state variable in each day are calculated using the state variables resulting from an initial model integration. These increments are then averaged through the month, and the state variables of the first day are updated using monthly averaged increments. The second run then propagates the updates into the remaining days of the month. Girotto et al. (2016) compared the proposed structure against two previously mentioned methods and concluded that the proposed method better reflects the submonthly variation in the TWS and its components. Further, two other approaches have been employed in other experiments. van Dijk et al. (2014) proposed a DA structure to merge the GRACE TWS retrievals and off-line products of several hydrological models on a monthly time step. Tian et al. (2017) employed an EnKS approach to jointly assimilate the GRACE TWS retrievals and SMOS-derived soil moisture into the Australian Water Resources Assessment-Land (AWRA-L) model. A notable common point in all published studies is the challenge of correctly assimilating the GRACE TWS retrievals, which are TWS anomalies. These anomalies cannot be directly assimilated into the model because it is not possible to predict the TWS anomalies without having a baseline (section 3.3.3). Therefore, in the above mentioned studies, these anomalies were converted to absolute values by adding a TWS baseline prior to the assimilation.

Although a number of different customized EnKF structures have been developed to assimilate GRACE TWS retrievals, a comprehensive comparison has not yet been conducted. Therefore, the first objective of this study was to compare these structures in a synthetic problem. To do so, two paradigms could be used. The structures of EnKFs can be considered as an integrated approach which cannot be decomposed. Then, these methods can simply be compared based on their performances. Conversely, a more complicated strategy is to consider these structures as modular frameworks. In this approach, each component of the structures could be analyzed separately. This leads to a better understanding of the structures. Moreover, as each structure has its own strengths and weaknesses, using a modular comparison provides an opportunity to improve the performances of the structures by combining components from different structures to achieve a more sophisticated EnKF structure. The latter paradigm helps to define a generic and modular EnKF framework with a number

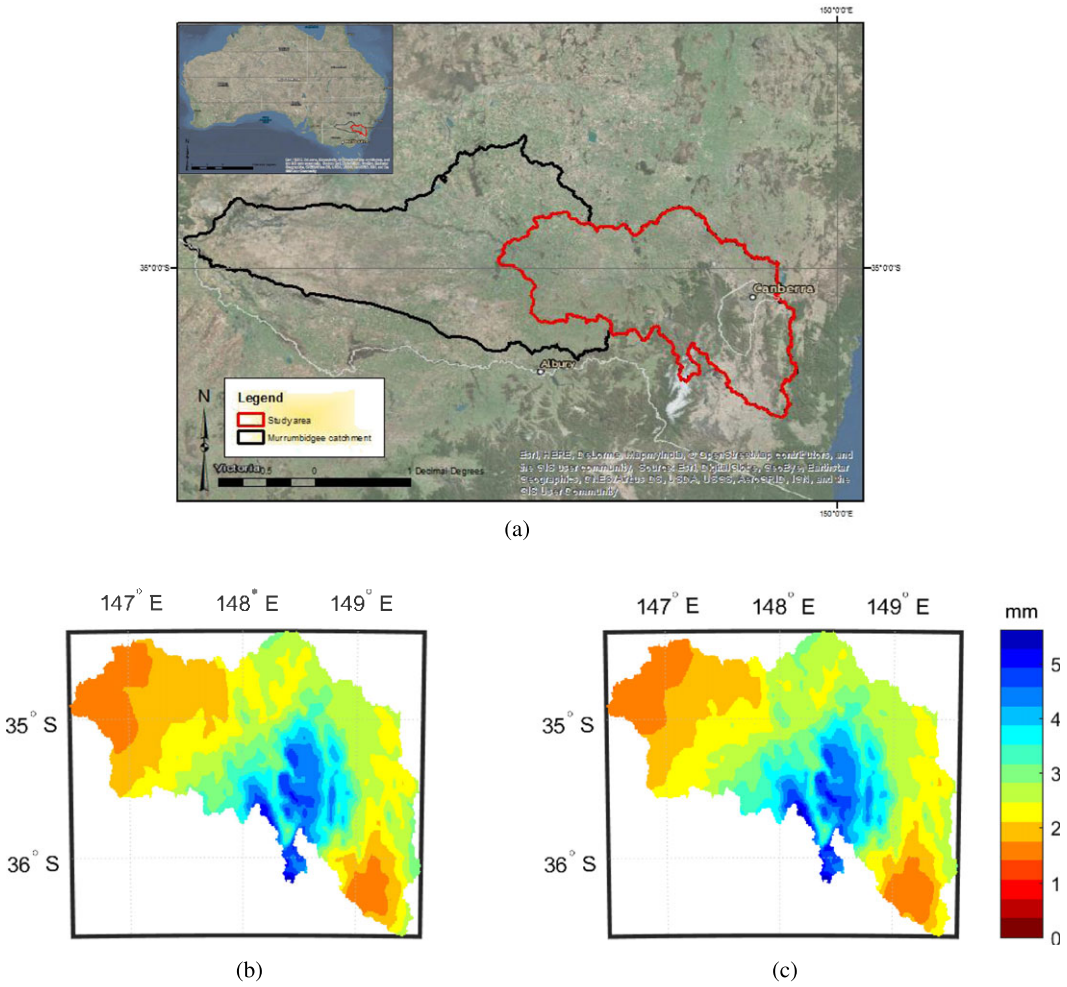


Figure 1. The top panel is the (a) study area map and its location, and the bottom panels are temporally averaged of unperturbed daily precipitation for (b) the open loop and data assimilation and (c) *truth* over a 4-year period (2007 to 2010).

of different alternative for each module. Instead of developing new structures, this framework, in addition to providing a sufficient understanding of the characteristics of each module, enables us to extract the optimum EnKF structure matching the modeler interests. In the current study, the second paradigm was used to compare the different EnKF structures. In general, the differences among the TWS assimilation approaches from previous studies are limited to two major components of the EnKF, the observation operator (mapping the modeled state variables onto GRACE observation space) and the Increment Calculation and Updating Strategy (ICUS) including the state variable selection approach and the increment implementing method. Therefore, these two components of the EnKF were considered as modules of the generic EnKF structure in this study.

Moreover, as was discussed, the GRACE anomalies were converted to absolute values by adding a baseline TWS. This baseline is mostly estimated based on a single baseline run with the assumption of an unbiased model. However, this assumption could be questionable in many cases so an alternative solution has been sought. Therefore, the second objective of this study was an alternative approach to assimilating the TWS anomalies without conversion to the absolute values. Instead, the TWS changes were used as the observation source. The TWS anomalies are the TWS absolute values minus the baseline, and the TWS changes are the average TWS absolute value of one month minus the average TWS absolute value of the previous month (more details are provided in section 3.3.3). A synthetic problem was used to compare the different approaches. The main reason for using a synthetic problem was the fact that the correct state variables and their uncertainty are known and controllable. Therefore, it provides a better insight into the opportunities and obstacles of the different approaches.

To address the study objectives, different combinations of the DA structures were tested to assimilate synthetic GRACE observations into a spatial hydrological model (AWRA-L). Because of the large number of assimilation structures that have been used, it is not clear which method provides the best results. This paper aims to provide an answer to this question. The study was performed in the upper Murrumbidgee catchment with a spatial resolution of 0.01°. The performance was evaluated by comparing the model results to a synthetic truth from an off-line model run.

2. Study Area and Model Description

The eastern half of Murrumbidgee catchment has been selected as the study area because of the availability of a large number of gauged subcatchments and the relative abscence of regulated streams (unlike the western part). This catchment is a part of the Murray Darling basin, located in southeastern Australia and covering approximately 3.5×10^5 km^2 of land (Figure 1a). The region is mostly mountainous with elevations ranging from over 2,250 m in the eastern ranges to approximately 250 m with a prevailing east to west slope. The precipitation ranges from 500 to 1,400 mm annually (Green et al., 2011).

The AWRA-L version 0.5, which is part of the AWRA system (van Dijk, 2010), has been selected as the hydrological model for the current study. The AWRA-L is a grid-distributed model simulating the streamflow, surface, shallow and deep soil moisture, and groundwater and vegetation water contents. This model has relatively few parameters compared to fully process-based spatially distributed hydrological models (which can have dozens of parameters per grid cell) and has been proven to work well in Australia (van Dijk et al., 2011, 2012; Renzullo et al., 2014).

In the AWRA-L model, two hydrological response units for each pixel are defined by default dividing the soil water and heat flux processes into two parts representing shallow and deep rooted vegetation. Conversely, the groundwater and surface water fluxes are simulated in an integrated process per pixel. It should also be noted that no lateral water distribution is considered between pixels. Water states are simulated in a number of different storages including surface soil moisture (S0), shallow soil moisture (Ss), deep soil moisture (Sd), groundwater (Sg), runoff (Sr), snow (Ssn), and vegetation (Sv) water content. The vertical water distribution among the state variables is controlled by the fluxes including infiltration, soil evaporation, drainage, and root water uptake. The AWRA-L daily temporal resolution is fixed, but its spacial resolution is adjustable. A 0.01° spatial resolution is selected for the current study.

The model inputs for this study are climate data (minimum and maximum temperature, precipitation, and radiation) from the Terrestrial Ecosystem Research Network Ecosystem Modelling and Scaling Infrastructure at 0.01° resolution. The wind speed was obtained from the Commonwealth Scientific and Industrial Research Organisation at 0.01° resolution (McVicar, 2011). The land cover map (Lymburner & Australia, 2011) was originally available with a 250-m resolution and upscaled to a 0.01° resolution. The model was calibrated using the observed streamflow data of 24 gauge stations from 2009 to 2010 (available online at http://realtimedata.water.nsw.gov.au/water.stm) using the Patient Rule Induction Method Parameter Estimation (PRIM-PE; Shokri et al., 2017). This method has the capability of identifying regions of the parameter space containing the parameter sets with *good-enough* (i.e., better than a threshold) performances. Using the PRIM-PE, all good-enough parameter sets were detected. However, since the model was not calibrated for the TWS components (including soil moisture and groundwater), different parameter sets could provide different estimates of the TWS. In the current study, to mimic the differences between real-world processes and the model, two good-enough parameter sets were employed. The first was used for the generation of the synthetic truth and the second for the DA. It should be noted that since the connection between groundwater and deep soil moisture was not calibrated, the model could not distinguish these two stores. Therefore, the deep soil moisture and groundwater were considered as a single store.

3. Methodology

3.1. Synthetic Experiment

The main objective of this study is to find an optimal EnKF structure for assimilating the GRACE TWS retrievals into a fine resolution hydrological model. To do so, the key structural features (hereafter referred to as structural variables; detailed description is provided in section 3.3) of the EnKF were first defined. All different possible EnKF structures formed by combining the structural variable alternatives were then examined in a synthetic setting.

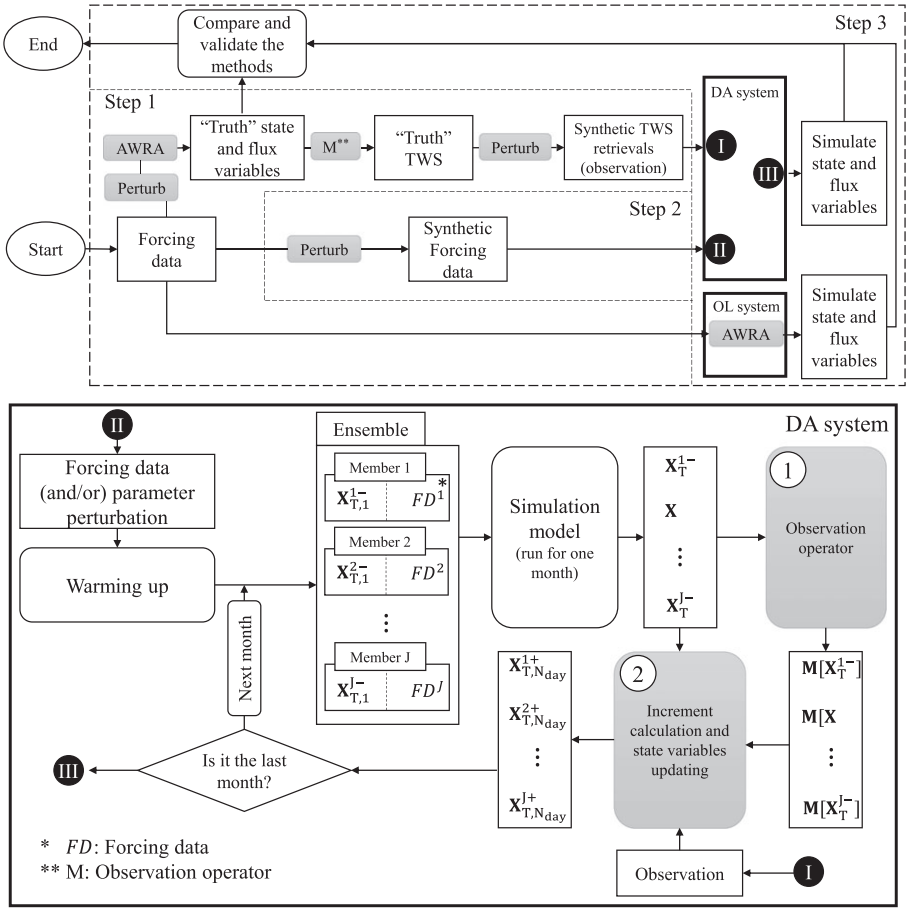


Figure 2. Schematic of the synthetic study. The top panel is the structure of the synthetic problem consisting of three steps: step (1) generating the synthetic observation, step (2) generating the perturbed forcing data, and step (3) using different DAs or OL integrations to produce state and flux variables. The bottom panel is the *DA system* part of top panel with more details. AWRA = Australian Water Resources Assessment; TWS = terrestrial water storage; DA = data assimilation; OL = open loop.

The design of the synthetic problem in this study is similar to the twin experiment in Forman and Reichle (2013). Figure 2 shows its conceptual framework which consists of three main steps. As a first step, a reference forcing data at 0.01° spatial and daily temporal resolution were generated. These were used to generate the synthetic TWS retrievals, with a monthly time step and a single value for the entire catchment. This process was started by adding white noise to the precipitation, minimum and maximum temperature, and radiation maps to generate the reference forcing data reflecting the errors in the forcing data estimation process. Those errors are temporally and spatially autocorrelated and cross-correlated, so the noise fields were generated likewise (see section 3.5). A truth estimate of the hydrological flux and state variables (e.g., soil water content) was then produced by forcing AWRA-L with the true forcing data. This simulation used different parameters values from those used in the DA and open loop (OL) simulations. Both of these parameter sets were selected from a good-enough region in the parameter space (Shokri et al., 2017). By vertically integrating the state variables relating to different components of the TWS, a true high-resolution product of the TWS was generated. The synthetic coarse-resolution observations were produced by spatial averaging of the model results. The temporal average at days 5, 15, and 25 were then calculated, mimicking the GRACE observation system. Observation noise, representing the errors in the real GRACE TWS retrieval estimation process, was then added to these averages. The upscaling procedure for the true TWS synthesizing was adopted from Zaitchik et al. (2008). The observation noise was assumed to be temporally (Forman et al., 2012; Su et al., 2010; Zaitchik et al., 2008) uncorrelated. Based on the synthetic observation area ($\sim 3.5 \times 10^5$ km^2), the standard deviation of the noises was assumed to be 20 mm, which is identical to the standard deviation used in Zaitchik et al. (2008), Su et al. (2010), and Forman and Reichle (2013) with a similar resolution.

Table 1
Structural Variable List

Structural variables		Alternatives
Increment calculation and updating strategies (10 alternatives, excluding the ICUS 1.3 and ICUS 3.3)[a]	Increment calculations	ICUS 1.y (Zaitchik et al., 2008)
		ICUS 2.y (Su et al., 2010)
		ICUS 3.y (-)
		ICUS 4.y (Girotto et al., 2016)
	Updating strategies	ICUS × .1 (Girotto et al., 2016)
		ICUS × .2 (Zaitchik et al., 2008)
		ICUS × .3 (Tian et al., 2017)
Observation operator (four alternatives)		H1 (-)
		H2 (Su et al., 2010)
		H3 (Zaitchik et al., 2008)
		H4 (Tian et al., 2017)
Observation type (two alternatives)		TWS absolute values (all previous studies)
		TWS changes (-)
Previous absolute observation[b] (two alternatives)		PAO 1 (-)
		PAO 2 (-)
State variable status[b] (two alternatives)		SVS 1 (-)
		SVS 2 (-)

Note. A sample (the earliest) of using each structural variable is mentioned in front of that. The variables followed by a dash (-) for the first time were used in the current study. ICUS = Increments Calculation and Updating Strategy; TWS = terrestrial water storage; PAO = previous absolute value; SVS = state variable status. [a]The ICUS alternatives are formed as the combinations of two different subvariables. In total, from 12 possible combinations, 10 are feasible to be employed in the EnKF structure. [b]Just applicable to assimilation of TWS changes (and not TWS absolute values).

After generating the maps of true state and flux variables and the synthetic observations, the second step consisted of generating the ensemble members. Each member was generated using a set of perturbed forcing data that reflect the errors in estimation process. In this study, it was assumed that these errors were identified correctly and accurately. Therefore, a number of noise fields with the same characteristics as the noise field were employed to perturb the forcing data used to generate the true forcing data.

The third and final step established the assimilation and the OL runs using the perturbed forcing data and the original (unperturbed) forcing data, respectively. The resulting flux and state variable estimates were compared to the true state and flux variables. This comparison assesses the ability of each approaches to reproduce the true variables.

3.2. Validation

The validation was performed by comparing the results of the DA runs to the synthetic truth. To do so, the bias and root-mean-square error (RMSE) were calculated and reported. The bias was computed as the average of the pixel errors (OL and DA minus truth), and RMSE was calculated as the square root of average squared pixel errors at each time step. The correlation coefficient (R) was also analyzed. R is mathematically related to bias and RMSE, and therefore, results are not separately discussed (but they are provided in the supporting information). Four different sets of synthetic observations were generated, to avoid drawing erroneous conclusions caused by the randomness in the process. For the same reason, each DA structure was applied three times to each set of synthetic observations. Thus, in total, each DA structure was applied 12 times. The distribution of the performance metrics including mean, median, and range were then analyzed.

3.3. DA Framework

As explained in section 1, the overarching objective of this study was to compare the modules of different customized EnKFs for assimilating the GRACE TWS retrievals. In this study, by investigating the existing EnKFs, the common and noncommon parts were detected and a generic and modular EnKF framework was designed.

The bottom panel of Figure 2 shows a flowchart of the DA strategy applied. The start is a perturbation of the forcing data (for more detailed descriptions see sections 3.1 and 3.5) in which a number of ensemble members

were generated. Then, to generate realistic initial states for each member, the following model initialization strategy was used. The warm-up (spin-up) consisted of two steps. First, by repeating the simulation of entire period 10 times (from 2006 to 2010) using the unperturbed forcing data, a reasonable estimate of the initial state was obtained. Then, each ensemble member was run for a further warm-up period of 2 years (from 2006 to 2007) using the perturbed forcing data to obtain initial states with a sufficient spread. The initial state variables after the warm-up period are denoted by $X^{j-}_{T,1}$ for the jth member. The main assimilation cycle (from 2008 to 2010) was then started by propagating the ensemble members for 1 month. During this step the state variables including time series of water storages in each pixel were generated. These are denoted as X^{j-}_{T}. This state vector thus consists of the water storages for all days of the month and every model grid in the domain.

The observation operator was then used to convert the state variables to the observation prediction for each member. The observation operator is one of the structural variables that was examined in the current study. Therefore, different alternatives were tested (see section 3.3.1). The next step of the assimilation procedure was the update step, which includes the calculation and implementation of the increments to the predicted state variables (X^{j-}_{T}). As different approaches were applied for this step, in this study it was considered as another structural variable. Different alternatives for this structural variable are described in section 3.3.2.

Previous studies performed a prior transformation to convert the TWS anomalies to the TWS absolute values by adding an estimated baseline (in most cases a long-term OL simulation). Here the performance of a new approach was evaluated against the transformation approach. This new approach assimilated the TWS changes into the model instead of the TWS absolute values, to avoid possible errors from baseline estimation (see section 3.3.3 for more details). So the observation type (absolute or change) was the third structural variable. To find the best DA structure, the effect of all different combinations of these structural variables was investigated. Table 1 shows a list of these structural variables.

3.3.1. Observation Operators

Three different observation operators were found in the literature. These operators as well as a new one (as a benchmark) were recoded and redesigned to enable them to be used interchangeably in the generic EnKF structure.

The observation operators are used to map the state variables onto the observation space. Hydrological models usually calculate the components of the TWS as separate state variables in a number of spatiotemporal fine grids. Therefore, the observation operator for assimilation of the TWS has to aggregate and upscale the TWS components. For this purpose, the spatial upscaling was performed by averaging the TWS of the contributing pixels. For the temporal upscaling, four different approaches (here they are referred as H1 to H4) were used. Equation (1) is a generic formulation of these observation operators:

$$\mathbf{M}(\mathbf{X}^{j}_{T}) = \frac{1}{N_{\text{pixel}}} \frac{1}{|\mathbf{D}|} \sum_{k=1}^{N_{\text{pixel}}} \left(\sum_{d \in D} \text{TWS}^{j}_{T,d,k} \right), \tag{1}$$

where N_{pixel} is the number of pixels in the observation grid, $\mathbf{D}$ represents the set of days, and $|\mathbf{D}|$ is the number of days that are averaged. $\text{TWS}^{j}_{T,d,k}$ is the TWS of the kth model pixel at the dth day of the Tth monthly time window of the jth member.

In operator H1, the spatially averaged TWS of the first day was considered as the representative of the entire month TWS ($\mathbf{D} = \{1\}$; $\{.\}$ is set notation). We did not find any examples of this approach in literature. Operator H2 (adopted from Su et al., 2010) is similar to H1 but uses the TWS of the last day as the representative ($\mathbf{D} = \{N_{\text{day}}\}$). Conversely, operators H3 (adopted from Zaitchik et al., 2008) and H4 (adopted from Tian et al., 2017) use an average of multiple days instead of a single day. More specifically, H3 uses $\mathbf{D} = \{5, 15, 25\}$ and H4 uses $\mathbf{D} = \{1, 2, \ldots, N_{\text{day}}\}$. A list of the observation operator alternatives is provided in Table 1.

Each of these observation operators which has been applied in the previous studies has its own philosophy. H2 assumes that the GRACE TWS retrievals are produced once a month (the last day of the month). Then, following the regular EnKF structures, whenever a new observation becomes available, the state variables of exactly that day should be updated so H2 uses the last day of the month to predict the observation. One can argue about the assumption that the observations belong to the last day of the months. Therefore, H1 is used in this study by transforming this approach to another extreme format in which it is assumed that the GRACE TWS retrievals belong to the first day of the month. These approaches seem to be a bit unrealistic as the GRACE TWS retrievals, unlike most of the other satellite products, are not snapshots of the water states. Conversely,

the GRACE provides estimates of water state changes during a time step. However, even with this imperfection, H2 and H1 provide some computation simplicity in the estimation process. To overcome the drawback of the H2 and H1, H3 calculates the TWS as an average of 3 days to mimic approximately three overpasses of the GRACE during a month. Finally, H4 uses an average of TWS of all days of the month to optimally simulate the GRACE TWS retrievals using state variables of models to potentially increase the accuracy of the observation prediction, specifically when it is used along with a smoother to update all state variables during the month.

3.3.2. ICUS Alternatives

Previous studies used different ICUS. The ICUS is a part of the EnKF process which calculates the increments and adds them to the state variables. These strategies can be decomposed into two main sections including the increment calculation section and the section where increments are used to modify the state variables. Three approaches of increment calculation and four updating strategies were found in the literature. These were then recoded and redesigned to let them be used interchangeably in the EnKF framework. The following describes them. The general equation to calculate the increments in the EnKF framework is

$$\Delta \mathbf{x}_T^j = \mathbf{K}_T[\mathbf{y}_T^j - \mathbf{M}(\mathbf{X}_T^j)], \tag{2}$$

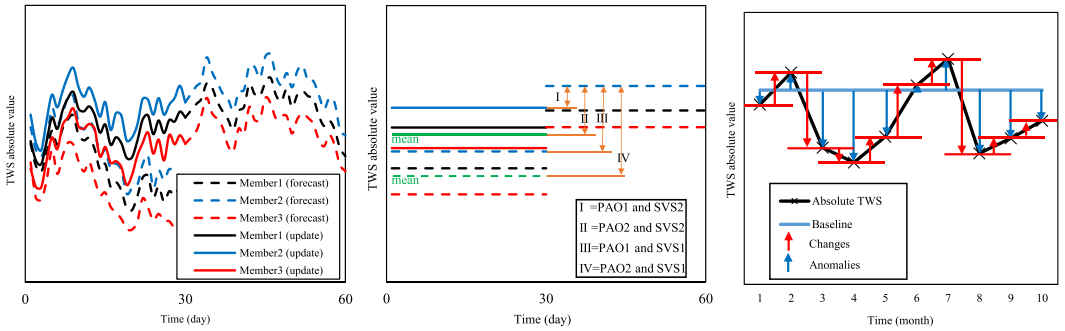
in which $\mathbf{y}_T^j$ is a vector that contains the perturbed observations, $\mathbf{M}(\mathbf{X}_T^j)$ is the simulated observation, and $\mathbf{K}_T^j$ is the Kalman gain. T refers to the assimilation time window, and j indicates the ensemble member. As can be seen in equation (2), the Kalman gain ($\mathbf{K}_T$) is a matrix which converts the innovations (difference between the observation and model prediction) into the state variable increments $\Delta \mathbf{x}_T^j$. It determines the relative weight of the observation and model during the update step by adjusting the magnitude of increments based on the error cross covariance between state variables and model observation predictions ($\mathbf{C_{xM}}$), the error covariance of the model observation prediction ($\mathbf{C_{MM}}$), and the observation error covariance ($\mathbf{C_{yy}}$). The formulation of the Kalman gain is

$$\mathbf{K}_T = \mathbf{C_{xM}}.(\mathbf{C_{MM}} + \mathbf{C_{yy}})^{-1}, \tag{3}$$

in which $\mathbf{C_{xM}}$ and $\mathbf{C_{MM}}$ are calculated based on the ensemble simulation results and $\mathbf{C_{yy}}$ reflects the uncertainty (or error) of the observations (Reichle et al., 2002). The $\mathbf{K}_T$ and consequently $\Delta \mathbf{x}_T^j$ can be calculated using all state variables or alternatively just a subset. In the latter case the increments to update the excluded state variables are not produced. This exclusion can occur in two situations: first, in case of some state variables are not observable and second, to simplify the assimilation procedure by reducing the number of state variables. In the TWS DA, the full state vector consists of the state variables of all pixels at all time steps within a month. Using this full state vector defines an approach that behaves similar to a smoother. However, theoretically, it is possible to assume that the increments of the state variables during a month are similar or can be estimated using just a subset. Therefore, it is possible to exclude some of the days from the increment calculation process and update the state variables just using the increment calculated from the remaining days. Recent studies used different subsets for this purpose. In some studies (e.g., Zaitchik et al., 2008), the increments were calculated just for a single time step, mostly the first or the last day of the month. However, in other studies (e.g., Girotto et al., 2016), the state variables of multiple days were employed to calculate the increments. Therefore, we tested four alternatives including the first day (ICUS1.y, adopted from Zaitchik et al., 2008); the last day (ICUS2.y, adopted from Su et al., 2010); days 5, 15, and 25 (ICUS 3.y); and all days of the month (ICUS4.y, adopted from Girotto et al., 2016).

After calculation of the increments, the next step was to add them to the state variables. In the previous studies, a number of approaches were proposed. As it is difficult to assign the GRACE TWS retrievals to a specific instance, one of the main disagreements among different approaches is about the time step(s) during the month which should be updated. Moreover, increments were applied to the state variables differently. Different strategies were adopted for this step and summarized in Table 1.

(1) In the first strategy, daily averaged increments were calculated and then added to the state variables of the first day of the month. Then, in order to obtain the updated value for the remaining time steps, the model was reinitialized and the current month resimulates using the updated initial conditions. This way, the calculated increments were propagated gradually through the month (ICUSx.1, adopted from Girotto et al., 2016).
(2) The second strategy is similar to the first but with the increments divided by the number of days and added to the state variables during the second run on a daily basis (ICUSx.2, adopted from Zaitchik et al., 2008).



(a) Modelled daily (spatially averaged) TWS.

(b) Temporally averaged TWS.

(c) TWS change and anomaly definition.

Figure 3. Schematic of the SVS and PAO structural variables. (a) Overview of the daily TWS absolute values from the forecast and update steps. (b) The temporally upscaled TWS for the forecast and update steps. The green lines indicate the TWS average values. As it is indicated in (b), to calculate the TWS change for Member 2 (blue), four approaches can be adopted. (c) The difference between the TWS anomaly and TWS change. SVS = state variable status; PAO = previous absolute value; TWS = terrestrial water storage.

(3) The third strategy is a direct approach, in which the increments that were calculated for each day were directly used just to update exactly that day (ICUS x.3, adopted from Tian et al., 2017). The advantage of this method is the fact that a resimulation is not required, which can save a considerable amount of computational effort. This implies that the state variables of the last day of the month have to be part of increment calculations. Hence, it can just be used with the second and fourth increment calculation strategy (ICUS2.3 and 4.3).

3.3.3. Observation Types

The observation operator role is to provide estimates of observations using the model state variable. The difference between the model prediction and observation (innovation) can then be used for the increment calculation. It should be noted that the observed values and the outputs of the observation operator should be the estimates of exactly the same variable. The GRACE retrievals are TWS anomalies (Figure 3c), so to directly assimilate them into a model, the observation operator must be able to predict the TWS anomalies using model state variables. Therefore, a TWS baseline is also required to convert the TWS absolute values to the TWS anomalies which are comparable with GRACE-based products (e.g., Schumacher et al., 2016, used this approach). Another alternative is to convert the observed TWS anomalies prior to the assimilation process using the TWS baseline (e.g., Forman et al., 2012; Girotto et al., 2017; Zaitchik et al., 2008, used this approach). It can be argued that both of these strategies are effectively the same.

As mentioned in section 1, the baseline in previous studies was estimated as an average of TWS in a long-term OL model integration. However, if the unbiased model assumption is not fulfilled, the use of this long-term average can lead to biased model results. To overcome this problem, a new approach is proposed here.

In this new approach, the observation operator was redesigned to generate the TWS changes (as the difference of TWS absolute values of two consecutive months) and the observed TWS anomalies (in this study synthetic TWS retrievals) were transformed to TWS changes. The changes were calculated between two consecutive months to make them consistent with the GRACE mission. In this way, both TWS absolute values (from model outputs) and the TWS anomalies (from observation) could be converted to estimates of exactly same value without the need for any external source of information (i.e., a baseline). This approach can be interpreted as partitioning each state variables into a fixed (the initial state at first day of the month) and a variable part (the change during the month and what the GRACE measures). Then, the variable part is updated using the TWS changes. The change in the synthetic TWS anomalies and the modeled TWS in the EnKF are then used as observation. Figure 3c shows the difference between the TWS changes and anomalies. The benefit of using the TWS changes is the fact that both the model-based TWS absolute estimates and the GRACE-based TWS anomalies can easily (without any external information) be converted to the TWS changes. The difference between these changes can then be used for increment calculation. However, the observation operator should be modified to reflect the difference between the TWS absolute values of the current and previous

time steps. The observation operator can be formulated as

$$\mathbf{M}_C(\mathbf{X}_T^j) = \mathbf{M}(\mathbf{X}_T^j) - \mathbf{M}_{T-1}, \tag{4}$$

in which $\mathbf{M}_C(\mathbf{X}_T^j)$ is the observation operator of the TWS change and $\mathbf{M}_{T-1}$ is the TWS absolute value estimate of the previous time step. In this equation $\mathbf{M}(\mathbf{X}_T^j)$ can be calculated from equation (4). However, multiple options exist to calculate $\mathbf{M}_{T-1}$. As shown in Figure 3b, different estimates of TWS absolute value are available. Either the TWS absolute value of the jth member from the forecast step, $\mathbf{X}_{T-1}^{-j}$ (dashed lines), or update step, $\mathbf{X}_{T-1}^{+j}$ (solid lines), can be considered as $\mathbf{M}_{T-1}$. Moreover, it could be argued that the average TWS estimates of all members, $(\sum_J^{j=1} \mathbf{X}_{T-1}^{+j})/J$ or $(\sum_J^{j=1} \mathbf{X}_{T-1}^{-j})/J$, might be a more accurate estimate of the previous TWS absolute values. Based on this description, for the observation type of TWS change two more structural variables are defined for assimilation of the TWS change, the state variable status (SVS) and the previous absolute value (PAO). The SVS1 and SVS2 indicate the use of state variables from the forecast and update step, while PAO1 and PAO2 indicate the use of the jth member and average of all members, respectively.

3.4. Experiment Setup

The different alternatives of the EnKF structure were formed as the combination of five structural variables. Considering 4 observation operators and 10 updating strategies for assimilating the TWS absolute values, 40 different structures were tested for the assimilation of the TWS absolute values. The number of alternatives for the case of the TWS change assimilation was 160, as four different approaches exist to convert the predicted absolute TWS to the change. All these experiments were conducted in two steps. First, the model was run for 24 months (from January 2006 to December 2007), after which the main DA process was applied to a 3-year period (from January 2008 to December 2010).

Since there were a number of random processes in the analysis, including the generation of the noise and perturbations in the EnKF, each structure was tested 12 times (4 different types observation noise × 3 trials for each); the standard deviation of the noise did not change in each of these repetitions; only the value of the random numbers was modified. The average of the performance metrics for each alternative was then considered as its total performance indicator. In all experiments, based on a sensitivity analysis, the ensemble size of 20 was selected. It was observed that by increasing the ensemble size to more than 20 members, the predicted observation (observation operator results) standard deviation and consequently the final results did not change significantly. This ensemble size is also consistent with previous studies on TWS assimilation. For example Zaitchik et al. (2008) used 20, Forman et al. (2012) used 16, and Girotto et al. (2016) used 24 members.

The observation for the current study was considered to be a single synthetic TWS retrieval for the entire catchment at a monthly time step. The full state vector ($\mathbf{X}_T^j$) consists of one surface water, one groundwater, and six subsurface storages (including surface, shallow, and deep soil moisture for two hydrological response units) per model grid per day, which leads to a vector size of 8,381,760 (for a 30-day month). However, as was mentioned previously, the state vector for different approaches differs in the number of days that were used in the increment calculation and updating step.

3.5. Generation of the Noise Fields

In order to generate a set of spatially and temporally autocorrelated and cross-correlated time series of the noise fields, an algorithm based on a discrete complex-valued spectral representation is used, which is implemented by the readily available fast Fourier transform. This algorithm has two main steps. First, the spatially autocorrelated and cross-correlated fields are generated, then the temporal correlation is added using a first-order autoregressive (AR (1)) process.

In the first step, generating J spatially autocorrelated and cross-correlated fields, each of the J random fields may be expressed as a Fourier-Stieltjes integral

$$q_j(x,y) = \int_{-\infty}^{+\infty} e^{i(\lambda x+\kappa y)} \mathrm{d}Z_{q_j}(\lambda,\kappa), \tag{5}$$

in which $q_j(x,y)$ is the value of jth random field at location x and y, i is $\sqrt{-1}$, λ and κ are the horizontal and vertical wave numbers. This equation can be written in a discrete form

$$q_j(x_n,y_m) = \sum_{l,p} e^{i(\lambda_l x_n+\kappa_p y_m)} \Delta Z_{q_j}(\lambda_l,\kappa_p), \tag{6}$$

Table 2
Perturbation Parameters

Parameter	Type[a]	Standard deviation	Spatial correlation	Temporal correlation	Cross correlation			
					Precipitation	Minimum temperature	Maximum temperature	Radiation
Precipitation	M	0.50	2°	3 days	n/a	−0.8	−0.2	−0.1
Radiation	M	0.30	2°	3 days	−0.8	n/a	0.6	0.5
Minimum temperature	A	0.30 (°C)	2°	3 days	−0.2	0.6	n/a	0.7
Maximum temperature	A	0.25 (°C)	2°	3 days	−0.1	0.5	0.7	n/a

[a]The symbols A and M indicate additive and multiplicative perturbations. n/a = not available.

in which, $\Delta Z_{q_j}(\lambda_l, \kappa_p)$ is defined by

$$\Delta Z_{q_j}(\lambda_l, \kappa_p) = \sqrt{\Delta\kappa_p \Delta\lambda_l} \times e^{\frac{\lambda_l^2+\kappa_p^2}{\sigma^2}} \times (\mathbf{c}_j.\mathbf{e}), \tag{7}$$

where σ is a parameter determining the spatial autocorrelation and

$$\mathbf{c}_j.\mathbf{e} = [c_{j1}, c_{j2}, \ldots, c_{jJ}].[e^{i\theta_{1,l,p}}, e^{i\theta_{2,l,p}}, \ldots, e^{i\theta_{J,l,p}}]^T, \tag{8}$$

with $\theta_{1,l,p}, \theta_{2,l,p}, \ldots, \theta_{J,l,p}$ uniform random uncorrelated fields between 0 and 2π that introduce random phase shifts and $c_{j1}, c_{j1}, \ldots c_{jJ}$ are the elements of a $J \times J$ matrix

$$\mathbf{C} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1J} \\ c_{21} & c_{22} & \cdots & c_{2J} \\ \cdots & \cdots & \ddots & \cdots \\ c_{J1} & c_{J2} & \cdots & c_{JJ} \end{bmatrix}, \tag{9}$$

After generating the spatially autocorrelated and cross-correlated noise fields, in the second step an AR (1) process was used to add the temporal autocorrelation according to

$$r_j(x_n, y_m, t) = \varphi_j \times r_j(x_n, y_m, t-1) + q_j(x_n, y_m), \tag{10}$$

in which φ_j adjusts the temporal autocorrelation with the correlation specifications given in Table 2. The autocorrelation settings are consistent with earlier studies (Forman et al., 2012; Girotto et al., 2016; Houborg et al., 2012; Zaitchik et al., 2008), and the error cross correlations are considered to be same as the cross correlation of the observed forcing data.

4. Results

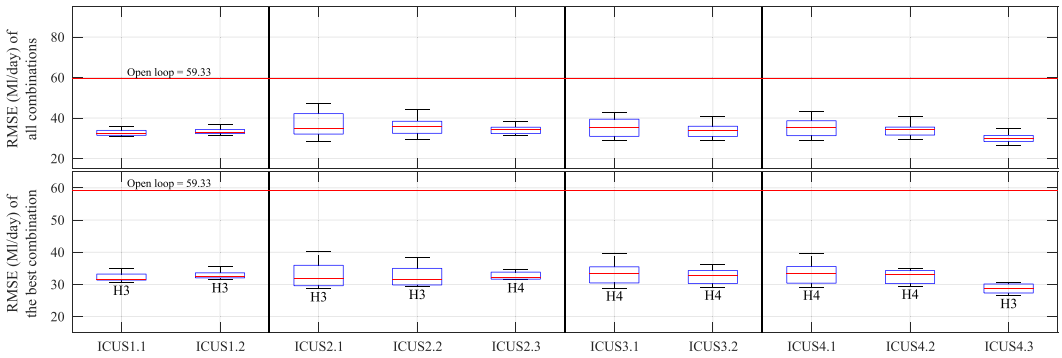
4.1. OL Results

The most influential meteorological forcing variable in the AWRA-L model is the precipitation (Renzullo et al., 2014). To provide a better understanding of the influence of precipitation uncertainty, the spatially averaged

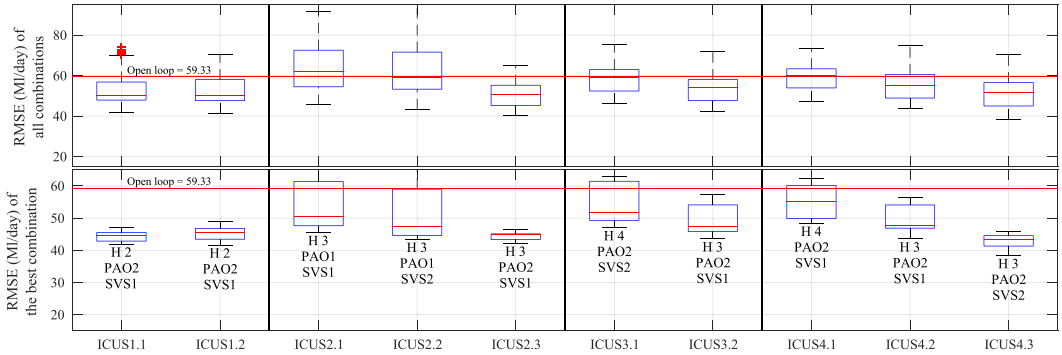
Table 3
RMSE, Bias, and Spatiotemporal-Averaged Values of the TWS and Its Components for the OL Run

Variable	RMSE (mm)	Bias (mm)	OL mean (mm)	Relative bias (bias/OL mean)
S0 (surface soil moisture)	2.60	−0.46	10.79	−0.042
Ss (shallow soil moisture)	14.83	−6.25	21.30	−0.293
Sd (deep soil moisture)	48.52	−34.36	85.98	−0.40
Sr (run-off)	0.98	−0.16	0.10	−1.60
TWS (terrestrial water storage)	59.33	−42.65	120.21	−35.48

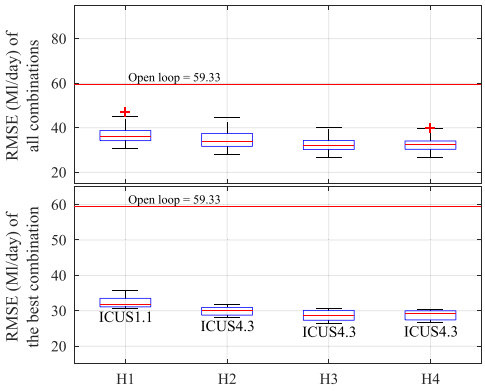
Note. RMSE = root-mean-square error; TWS = terrestrial water storage; OL = open loop.



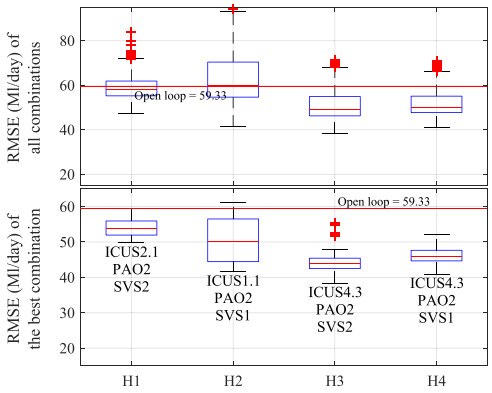
(a) Assimilation of the TWS absolute values. 48 values (4 H×12 trials) for each ICUS.



(b) Assimilation of the TWS changes. 92 values (4 H×2 PAO×2 SVS×12 trials) for each ICUS.



(c) Absolute TWS values. 120 values (10 H×12 trials) for each H.



(d) TWS changes. 480 values (10 H×2 PAO H×2 SVS×12 trials) for each H.

Figure 4. RMSE box plot of all the combinations and the best combination for each of the ICUS (a, b) and observation operators (c, d). For clarity, different scales were used for the vertical axes. ML/day = 10^6 L/day.
RMSE = root-mean-square error; ICUS = Increments Calculation and Updating Strategy; TWS = terrestrial water storage; PAO = previous absolute value; SVS = state variable status.

map of the unperturbed precipitation used to generate the synthetic truth and used in the OL and DA runs is depicted in Figures 1b and 1c. It can be seen that the OL and DA have less precipitation than the truth, which is expected to cause a negative overall bias in the estimate of the state variables, including the TWS. Furthermore, as the precipitation noise was added multiplicatively, a higher difference emerges in the central parts. Consequently, higher errors in the model outputs are expected for this region. It is expected that the model systematically underestimates the TWS, especially in the central region. This underestimation is not just limited to the TWS. As shown in Table 3, the values of all TWS components are lower for the OL run than for the truth. The highest negative bias and RMSE belong to the deep soil moisture, which had the highest storage capacity and consequently the highest contribution in the forming TWS.

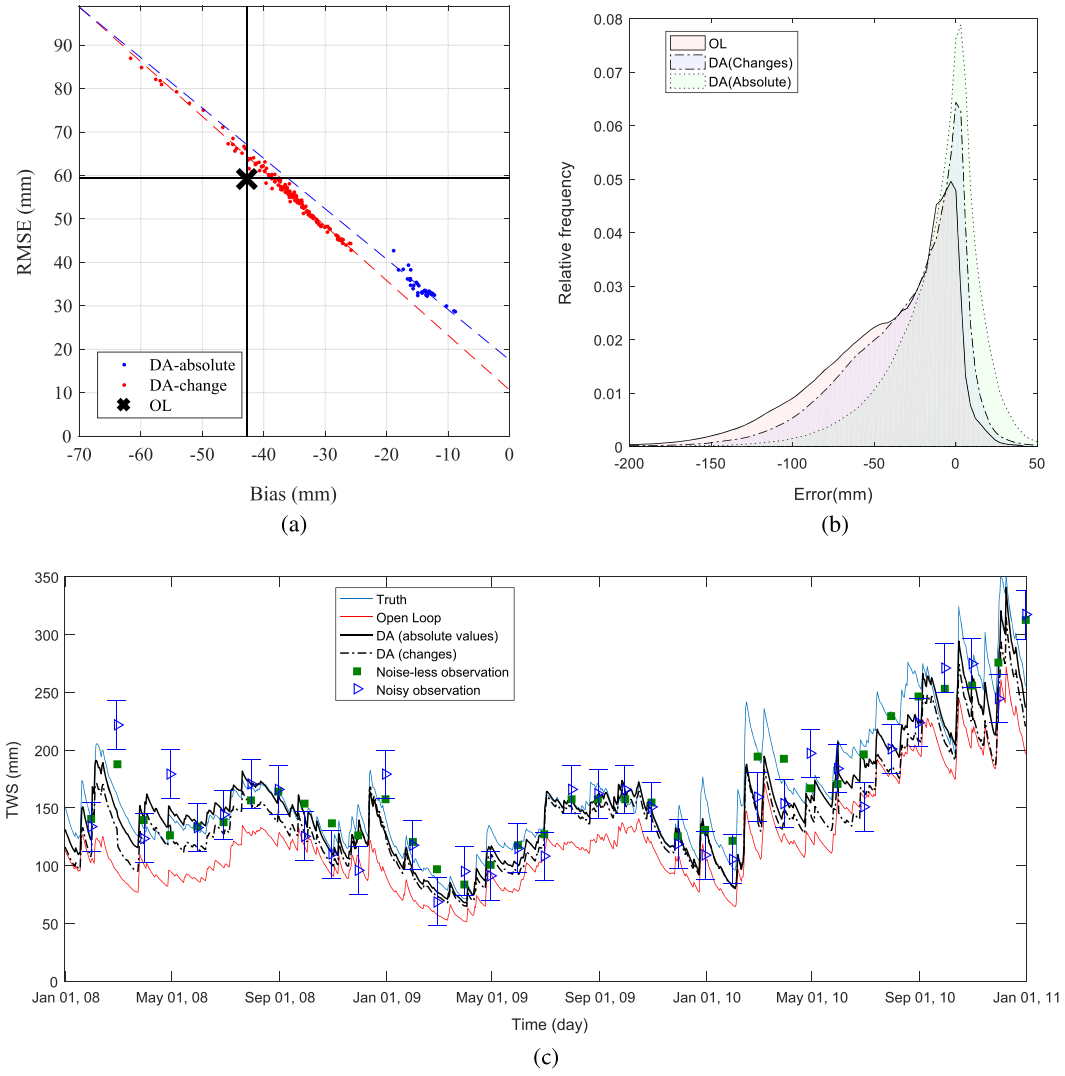


Figure 5. (a) Average performances of the different DA approaches when assimilating the ideal TWS absolute value (formed by adding an unbiased and accurate baseline to the TWS anomalies) and the TWS changes. (b) TWS error histogram from the OL and the best EnKF structures for TWS change assimilation (ICUS4.3, H3, PAO2, and SVS2) and TWS absolute value assimilation (ICUS4.3, and H3). (c) Results of the use of the best combination of assimilation structures, when assimilating TWS absolute values (ICUS4.3 and H3) or TWS changes (ICUS4.3, H3, PAO2, SVS2). The noise-free observation is the value of TWS absolute values before adding the noise in the observation synthesizing process. DA = data assimilation; TWS = terrestrial water storage; OL = open loop; EnKF = Ensemble Kalman filter; ICUS = Increments Calculation and Updating Strategy; PAO = previous absolute value; SVS = state variable status.

4.2. DA Results

To optimize the EnKF structure, five different structural variables were defined, as indicated in Table 1. However, it should be noted that the PAO and SVS are not applicable when assimilating the TWS absolute value. Therefore, in total, 40 feasible structures for the assimilation of the TWS absolute value and 160 (4×40) feasible structures for the assimilation of the TWS changes were tested. In this section, the effect of employing different options of the structural variables was analyzed.

4.2.1. ICUSs

As explained in section 3.3.2, 10 different ICUSs were tested. These were formed by two variables: (i) the time step(s) that was (were) used to calculate the increments and (ii) the updating strategies. In Figures 4a and 4b the top panels represent the RMSE values obtained from all trials of all possible combinations (16×12 values for each box plot in the TWS absolute value assimilation and 4×12 values for each box plot in the TWS change assimilation) that used each particular ICUS. The bottom panels show the results of the 12 trials for the best combination of the structural variables. This best structural variable is stated underneath the box plots.

When TWS absolute values were assimilated into the model, all structures performed better than the OL (Figure 4a). However, some of the ICUSs showed higher potential for reduction of the RMSE values. ICUS4.3 had the best potential to improve the performance of the EnKF (Figure 4a). This ICUS behaves more similar to a smoother. It means that the increment is calculated for each time step and then directly applied to the state variables of the same time step. The problem with this ICUS is its huge number of state variables and consequently the high computational demand of the assimilation procedure. Because the assimilation algorithm requires temporally uncorrelated observation errors and the smoother assimilates a single observation into the model during each month, Girotto et al. (2016) concluded that the application of smoothers in the GRACE assimilation is inappropriate. However, the results described here show that the ICUS4.3 provides the best results, at least for this synthetic problem.

The bottom panel of Figure 4a shows that, except for ICUS4.3, all ICUSs show a relatively similar potential to improve the model results. However, some ICUS provided more consistent results, as evidenced by the smaller spread in the results of the 12 trials for ICUS1.1, 1.2, and 2.1. Regarding the assimilation of the TWS changes (Figure 4b), although important information (the estimate of the baseline) was removed from the assimilation process, most of the ICUSs improved the performance in comparison with the OL. Even in the case of assimilation of TWS changes, ICUS4.3 led to the best performance, while ICUS1.1, 1.2, and 2.1 showed similar RMSE values. Moreover, it can be seen that using the best structures, the RMSE values were lower for the DA runs than the OL for all of the ICUSs. However, some of the ICUSs led to an increase in RMSE.

Based on the results of both the observation types (TWS changes and absolute values), those ICUSs that update the state variables using the increments from the same days had the highest reliabilities (i.e., ICUS4.3, 1.1, and 2.3). Except for these structures, ICUS1.2 was the only structure showing results similar to the best. However, its results were still slightly worse. This ICUS breaks down the increments of the first day and updates the state variables by adding on gradually during the second run.

For each structure, the RMSE and bias average of all 12 trials were used in this experiment (Figure 5a). When ideal estimates of the TWS absolute values (i.e., TWS absolute values which are estimated using an accurate baseline) exist, assimilating those is preferable to assimilating the TWS changes. However, without such ideal estimates of the TWS absolute values, assimilation of the TWS changes can still provide an acceptable outcome if the structure of the filter is set correctly. As can be seen in the figure, the trend line in the case of TWS changes assimilation is shifted toward the lower left as compared to the TWS absolute values trend line. This indicates that the TWS change assimilation systematically reduces the RMSE more than the bias. One reason for this outcome is that the assimilation of TWS absolute values the EnKF had information about the long-term TWS average which was not introduced to the TWS change assimilation. As this information is directly related to the bias, the EnKF with absolute values can remove the bias more efficiently.

4.2.2. Observation Operator

Another structural variable is the observation operator, which had four alternatives. Figures 4c and 4d show the box plot of the performance for several structures using each observation operator alternative for the assimilation of TWS absolute values (Figure 4c) and TWS changes (Figure 4d). In these figures the top panels represent the RMSE values obtained from all possible combinations (40×12 values for each box plot in the TWS absolute value assimilation and 10×12 values for each box plot in the TWS change assimilation) that used each particular observation operator. The bottom panels show the RMSE values of only the best combinations. All experiments showed a better performance than the OL, and RMSE was slightly lower when H3 and H4 were employed (Figure 4c). As mentioned in the figure, the best ICUS for the H2, H3, and H4 was ICUS4.3 and in the case of H1 the best performance was observed for a combination with ICUS1.1. The combination of H1 and ICUS4.3 had a very similar performance as the combination of H1 and ICUS1.1.

In the case of the TWS change assimilation a different pattern was observed. As depicted in Figure 4d the performance of the assimilation was more sensitive to the selected structure. Therefore, in this case it is crucial to correctly select the structure. As can be seen in the top panel, using only the best observation operator alternatives (i.e., H3 or H4) and not optimizing the other structural variables, one can expect to have results with improved performances. However, there is a chance of producing worse results. Conversely, a combination of the best observation operator and optimal other structural variables could ensure an improvement in the results. These optimum settings were a combination of ICUS4.3, PAO2, and either SVS1 or 2. Comparing the results obtained from H3 and H4 shows that they led to a similar performance as a result of their similar

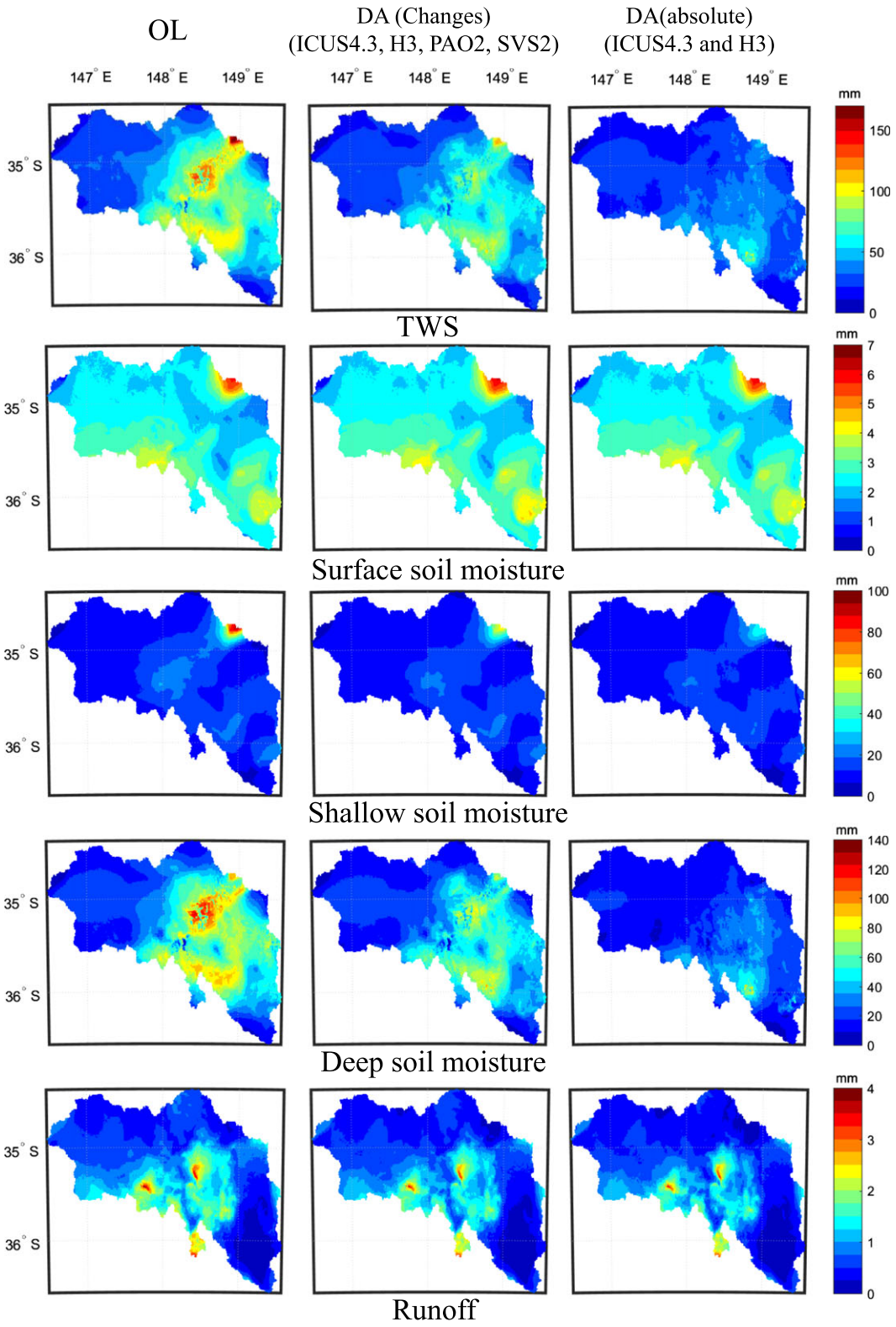


Figure 6. Temporally averaged RMSE of TWS, surface, shallow and deep soil moisture and runoff, from the OL and the selected DAs approaches using TWS changes (ICUS4.3 and H3) and TWS absolute values (ICUS4.3, H3, PAO2, and SVS2). RMSE = root-mean-square error; TWS = terrestrial water storage; OL = open loop; DA = data assimilation; ICUS = Increments Calculation and Updating Strategy; PAO = previous absolute value; SVS = state variable status.

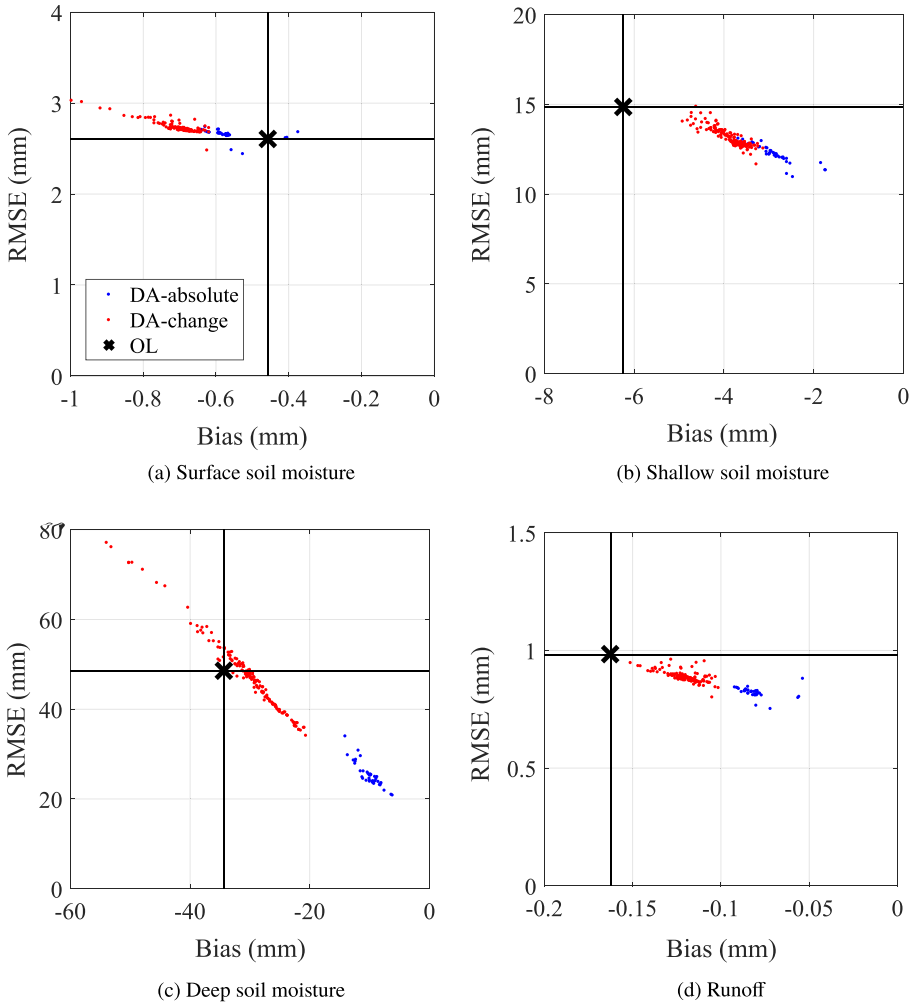


Figure 7. Impact of the assimilation of TWS absolute values or changes on the different TWS components, using the optimal assimilation strategy (combination of ICUS4.3 and H3 for TWS absolute value and combination of ICUS4.3, H3, PAO2, and SVS2 for TWS change assimilation).

structures. However, H3 provided a slightly wider range of RMSE values. This means that although using H3 provided a better RMSE than H4 on average, there is a chance of getting worse results.

When applying the optimal structure, all observation operators led to an improvement in the model results. In the case of using H2 and optimized other structural variables, the expected RMSE values were reduced as compared to the OL, and in the best case the RMSE was similar to the results obtained with H3 and H4. However, its performance was more uncertain compared to the performance of H3 and H4. This means that one can expect to gain some improvement in the results when using H2, but this improvement will not be as consistent as when using H3 and H4. The expected improvement using H1 along with an optimized set of other structural variables was the worst among the H alternatives. However, it was still better than the OL. Contrary to when using H2, even the best model performance was not comparable to the results when using H3 and H4. It can be concluded that the best observation operators were either the H3 or H4. As both of these observation operators predict the TWS as an average of multiple days, their resulting observation predictions and consequently their effect on the final products were similar. With a closer look at the RMSE values, it was observed that the differences between the performances of the H3 and H4 were not significant (i.e., less than 0.2-mm difference in the average RMSE).

Among the four best ICUSs, the best combination using ICUS1.1 and ICUS1.2 was formed by H2, PAO2, and SVS1 and the best combination of ICUS2.1 and ICUS4.3 was formed by H3, PAO2, and either SVS1 or 2 (as mentioned previously, the SVS did not have a significant effect on the results). This means that for the

assimilation of the TWS changes, if the increments are supposed to be calculated just for the first day of the month (ICUS1.X) the best observation operator is H2 (consider the TWS of the last day of the month as the representative of the month) so that the TWS changes are calculated as the difference between the last day of the previous month and that of the current month. Moreover, the results showed that PAO2 is preferable to PAO1. It means that in order to convert the TWS absolute values into TWS changes, the best approach is to use the average of all members' TWS. The low sensitivity of the results to the SVS also indicates that neither of the state variable sets, from the forecast or update step, was superior to another.

4.2.3. Results of the Best Combinations

Some detailed results of the best DA approaches are presented and discussed in this section. Figure 5c shows the spatially averaged truth and simulated TWS from the OL and the best DA approaches when assimilating the TWS absolute values (top panel) and TWS changes (bottom panel). In this figure, the whiskers represent the observation standard deviation. As shown in Figure 5c, for the current experiment the DA can correct the errors and this correction was more effective when ideal absolute TWS observations were assimilated. In the case of assimilating the absolute values, the DA TWS estimates were in general between the observations and OL TWS estimates and sufficiently close to the truth. In the case of DA using TWS changes, mostly the TWS estimates were improved. However, in only a few instances (i.e., near May 2010) the DA estimates became worse than the OL ones. Another point that should be noted is the larger errors between the simulations (DAs and OL) and the truth near the peaks. This can be explained by the model's tendency to overestimate the drainage during wet periods. However, the results show that the assimilation procedure can partly correct these errors.

The top panel of Figure 6 investigates the spatial performance of the selected DAs. This figure shows the temporally averaged RMSE of the TWS for the OL, the best DA strategy for assimilating the TWS changes and the best DA strategy for assimilating the TWS absolute values. As the top panel of Figure 6 shows, the averaged RMSE at the central area is significantly higher than for other regions of the study area. The main reason is the higher precipitation and consequently the higher variability of the perturbations in that region. As can be seen in the top panel of Figure 6, the errors were partially detected and removed by the DA methods. Moreover, the accuracy of the simulation in the eastern and western area was also improved through the DA. To provide a better understanding of the TWS estimate errors using the optimum DAs, the TWS error histograms are shown in Figure 5b. These errors are the differences between the simulated (OL and DA) and the truth per pixel and time step. The errors of the OL estimates are strongly left skewed, which indicates the OL negative bias. By applying the optimum DA, both the skewness and dispersion of the OL estimate were reduced but this reduction was less significant in the case of assimilating the changes.

4.2.4. Vertical Disaggregation

One of the objectives of the TWS assimilation is to update the storage values of all the different TWS compartments. Figure 7 shows the scatter plots of the bias against the RMSE of the DA approaches and the OL for the different components of the TWS including: surface soil moisture, shallow soil moisture, deep soil moisture, and run-off. Figure 7 shows that the most significant improvement occurred in the deep soil moisture followed by the shallow soil moisture. Run-off was slightly improved by the DA, while the surface soil moisture estimates were slightly degraded. A spatial illustration of these differences in the accuracy of the estimates can be seen in Figure 6. Among these components, the surface soil moisture estimates were either not improved or degraded (i.e., bottom left of the middle map). The shallow soil moisture (second row) and runoff (fourth row) estimates were slightly improved mostly in the region with a higher degree of error. By up to 52% and 80% correction in the RMSE and bias compared to the baseline simulation, the estimates of the deep soil moisture improved the most among all other components. The main reason of these differences relates to the contribution of each storage in forming the TWS. As represented in Table 3, the deep soil moisture had the strongest contribution, followed by shallow soil moisture. This finding justifies the approach of Tian et al. (2017) in jointly assimilating SMOS-derived surface soil moisture and GRACE-TWS to compensate for the inability of GRACE to improve the surface soil moisture.

5. Discussion

To determine an optimal structure of the EnKF to assimilate the GRACE coarse TWS retrievals, five different structural variables were identified and evaluated. For this purpose, a generic EnKF framework, with capability of implementing different structural variables, was developed. Using a synthetic experiment, the different EnKF structures were compared. The structural variables were the ICUS (10 alternatives), observation

operator (4 alternatives), observation type (2 alternatives), PAO (2 alternatives), and SVS (2 alternatives). In total, 200 feasible structures (160 when assimilating TWS changes and 40 when assimilating absolute values) were tested. The impact of employing the different options were subsequently analyzed. The reason for the smaller number of TWS absolute value tested is the fact that the PAO and SVS were not applicable in this case.

This study has a number of limitations. First, we used an off-line model forced with unperturbed meteorological data as synthetic truth. In reality, a considerable amount of uncertainty drives from the model structure. This uncertainty cannot be fully replicated when the same model structure is used for DA/OL. In our experiment, this shortcoming was partially mitigated by using two independent good-enough parameter sets found by the PRIM-PE method. Second, our study focused on a small subbasin in Australia without significant snowfall and used a 3-year simulation period. Additional experiments would be required to determine whether the different structures show similar differences in performance for other regions, period duration, resolution, or simulation models. We argue that the comparisons made here are still valid as all the structures were compared for an identical situation (even identical random numbers). In this study, a synthetic problem was used to compare different EnKF structures. While those results were informative, it should be noted that there is the possibility that the DA results change marginally with structure in a real-world application due to different conditions and/or the assumptions of this study. However, considering (1) the fact that the model was calibrated using real observations and used real input data and (2) the strong improvement in the performances of the best DA structures, the conclusions are expected to remain unchanged when implemented in a real-world scenario.

Currently, different GRACE TWS solutions using different gravity models and resolutions have been developed. In the current study the observations were considered to be similar to the official GRACE mission solutions. However, other solutions differ in terms of error correlation and resolution and the impact of these could be investigated separately.

When an ideal unbiased absolute TWS was used as the observation source, all the DA structures could improve the TWS estimate in comparison to the OL estimate. By using different DA approaches, bias and RMSE were improved in the range of 55–80% and 40–60%. This shows that, although all the structures improved the estimates, there was a considerable difference between different structures. Therefore, selecting an optimal DA structure is beneficial. As for the current problem among all tested structures the combination of ICUS4.3 and H3 had the best performance. This combination suggests that the best estimates for this case can be obtained using an approach which behaves more similarly to a smoother and observation operator which estimates the observation as the average of multiple days (5, 15, and 25). H4, as another observation operator that averages the TWS of all days in the month, had a very similar result to H3. This can be explained by the fact that in both H3 and H4, a number of days contribute similarly to the estimation process. It is also observed that among ICUSs, those that update the state variables of different days using the increments from the exact same days had the highest reliabilities in producing accurate estimates. Except for these structures, ICUS1.2 (which breaks down the increments of the first day and updates the state variables by adding it gradually during the second run) was the only structure that had acceptable results.

Assimilating the TWS changes rather than absolute values of the TWS was proposed as an alternative solution when an unbiased estimate of absolute TWS is not available. Using this approach, most DA structures still outperformed the OL, but their results were not better than those from the assimilation of the ideal absolute values. If an unbiased estimate of the absolute TWS exists, assimilation of the absolute values leads to a better performance. Otherwise, there are still benefits in assimilating the changes. When the TWS changes were assimilated, among the observation operators, H3 and H4 had the best performances, and the best ICUS was ICUS4.3. The combination of ICUS4.3, H3, PAO2, and SVS2 showed the best performances with around 40% and 28% improvement in the average bias and RMSE. Another finding regarding the assimilation of TWS change is that, if the increments are supposed to be calculated just for the first day of the month (ICUS1.X) the best observation operator is H2 (consider the TWS of the last day of the month as the representative of the month). In this approach the TWS changes were calculated as the difference between the last day of the previous month and that of the current month. Moreover, in order to convert the estimation of the TWS absolute values into TWS changes, using the average of all members' TWS as the estimate of the previous month TWS led to more accurate results. Also, the low sensitivity of the results to the SVS indicates that neither of the state variable sets, from forecast or update step, is superior to another. Assimilating the TWS changes may seem unrealistic (which is why the assimilation of TWS absolute values outperforms it). However, it still

improve overall estimation results. This is where a trade-off emerges between accepting the error from adding an estimated baseline or using the TWS change assimilation approach.

The largest impact of assimilation was on the largest storages. In this case the deep and shallow soil moisture were improved the most. Also, some improvements were observed in the runoff estimates. The surface soil moisture was the only state variable that was degraded slightly by all the DA approaches. This justifies the approach of Tian et al. (2017) in jointly assimilating the surface soil moisture and the TWS.

Beside the accuracy, a key characteristic of the DA approaches is their computational burden. This becomes more important when the spatial resolution of the model and/or the number of state variables per pixel increases. The DA structures that were applied in the current study had different computational complexities, and the main controller structural variable was the ICUS. By increasing the number of days that are used in the increment calculation procedure, the memory demand increases. From this perspective, the ICUS that used just one day was the most computationally efficient. Regarding ease of implementation, the ICUS with the third strategy of increments implementation (ICUSx.3) was the easiest to implement, because they need least effort to apply the filter, as there is no need for a second run and reinitializing.

6. Conclusions

This study undertook a synthetic experiment to assist in designing an optimal GRACE TWS DA approach more insightfully, by gathering and comparing all existing EnKF structures to assimilate the GRACE TWS into the land surface models. The results suggested that all structure could improve the results when absolute TWS was available, but assimilating the changes is an alternative that can retrieve the improvement partially even without having an accurate baseline to convert the GRACE TWS anomalies to accurate TWS absolute values. Moreover, results showed that even using identical observations, the performances of different EnKF structures considerably vary. Therefore, optimizing the EnKF structure is beneficial for assimilating the GRACE TWS. Assimilating the TWS had a significantly larger impact on the storages with a greater contribution to the TWS. Therefore, to improve the storages with low capacity (e.g., surface soil moisture), assimilating the TWS was not sufficient.

This study could be followed up in a number of different ways. Even though they were outside the scope of this study. Even though the synthetic problem led to a comprehensive insight from different points of view into the EnKF structures, it imposes some sources of uncertainties. Partially verifying these results using available data in a real-world problem should be considered in future studies. Another limitation of this study was that the study was performed for a specific region and simulation model. Therefore, other regions and simulation models can be investigated in the future.

Acknowledgments

Ashkan Shokri acknowledges the support from Monash University in the form of a Monash Graduate Scholarship (MGS) and Monash International Postgraduate Research Scholarship (MIPRS). Valentijn Pauwels is supported through an ARC Future Fellowship grant (FT130100545). The research is supported through an ARC Discovery grant (DP140103679). This research is also supported in part by the Monash eResearch Centre and eSolutions-Research Support Services through the use of the Monash Campus HPC Cluster. Climate data (minimum and maximum temperature, precipitation, and radiation) are obtained from the Terrestrial Ecosystem Research Network (TERN) Ecosystem Modelling and Scaling Infrastructure (eMAST) (available online at http://portal.tern.org.au/precipitation-driest-month-1970-2012/20317). The wind speed is obtained from the Commonwealth Scientific and Industrial Research Organisation (CSIRO; McVicar, 2011; DOI: 10.4225/08/5a6ffda5bb332). The land cover map is obtained from Geoscience Australia (Lymburner & Australia, 2011; available online at http://pid.geoscience.gov.au/dataset/ga/70138). The observed streamflow data are obtained from the New South Wales (NSW) department of industry-lands and water (available online at https://realtimedata.waternsw.com.au/water.stm).

References

Aubert, D., Loumagne, C., & Oudin, L. (2003). Sequential assimilation of soil moisture and streamflow data in a conceptual rainfall-runoff model. *Journal of Hydrology*, *280*, 145–161.

Barrett, A. P. (2003). National operational hydrologic remote sensing center snow data assimilation system (SNODAS) products at NSIDC. National Snow and Ice Data Center, Cooperative Institute for Research in Environmental Sciences.

Clark, M. P., Rupp, D. E., Woods, R. A., Zheng, X., Ibbitt, R. P., Slater, A. G., et al. (2008). Hydrological data assimilation with the ensemble Kalman filter: Use of streamflow observations to update states in a distributed hydrological model. *Advances in Water Resources*, *31*, 1309–1324.

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*, 1–16.

De Lannoy, G. J. M., & Reichle, R. H. (2016). Global assimilation of multiangle and multipolarization SMOS brightness temperature observations into the GEOS-5 catchment land surface model for soil moisture estimation. *Journal of Hydrometeorology*, *17*, 669–691.

Dechant, C., & Moradkhani, H. (2011). Radiance data assimilation for operational snow and streamflow forecasting. *Advances in Water Resources*, *34*, 351–364.

Eicker, A., Schumacher, M., Kusche, J., Döll, P., & Schmied, H. M. (2014). Calibration/data assimilation approach for integrating GRACE data into the watergap global hydrology model (WGHM) using an ensemble Kalman filter: First results. *Surveys in Geophysics*, *35*, 1285–1309.

Ellett, K. M., Walker, J. P., Western, A. W., & Rodell, M. (2006). A framework for assessing the potential of remote-sensed gravity to provide new insight on the hydrology of the Murray-Darling basin. *Australasian Journal of Water Resources*, *10*, 125–138. https://doi.org/10.1080/13241583.2006.11465286

Entekhabi, D., Rodriguez-Iturbe, I., & Castelli, F. (1996). Mutual interaction of soil moisture state and atmospheric processes. *Journal of Hydrology*, *184*, 3–17.

Evensen, G. (2003). The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean Dynamics*, *53*, 343–367. https://doi.org/10.1007/s10236-003-0036-9

Forman, B. A., & Reichle, R. (2013). The spatial scale of model errors and assimilated retrievals in a terrestrial water storage assimilation system. *Water Resources Research*, *49*, 7457–7468. https://doi.org/10.1002/2012WR012885

Forman, B. A., Reichle, R. H., & Rodell, M. (2012). Assimilation of terrestrial water storage from GRACE in a snow-dominated basin. *Water Resources Research*, *48*, W01507. https://doi.org/10.1029/2011WR011239

Girotto, M., De Lannoy, G. J. M., Reichle, R. H., & Rodell, M. (2016). Assimilation of gridded terrestrial water storage observations from GRACE into a land surface model. *Water Resources Research*, *52*, 4164–4183. https://doi.org/10.1002/2015WR018417

Girotto, M., De Lannoy, G. J., Reichle, R. H., Rodell, M., Draper, C., Bhanja, S. N., & Mukherjee, A. (2017). Benefits and pitfalls of GRACE data assimilation: A case study of terrestrial water storage depletion in india. *Geophysical Research Letters*, *44*, 4107–4115. https://doi.org/10.1002/2017GL072994

Green, D., Petrovic, J., Moss, P., & Burrell, M. (2011). *Water Resources and Management Overview: Murrumbidgee Catchment*. Sydney: NSW Office of Water.

Hendricks Franssen, H. J., Zhang, H., Kurtz, W., Kollet, S., & Vereecken, H. (2017). Improved characterization of root zone soil moisture in land surface models by assimilation of groundwater level data. An example with TerrSysMP. In *EGU General Assembly Conference Abstracts* (pp. 15653).

Hirschi, M., Mueller, B., Dorigo, W., & Seneviratne, S. I. (2014). Using remotely sensed soil moisture for land–atmosphere coupling diagnostics: The role of surface vs. root-zone soil moisture variability. *Remote Sensing of Environment*, *154*, 246–252.

Houborg, R., Rodell, M., Li, B., Reichle, R., & Zaitchik, B. F. (2012). Drought indicators based on model-assimilated gravity recovery and climate experiment (GRACE) terrestrial water storage observations. *Water Resources Research*, *48*, W07525. https://doi.org/10.1029/2011WR011291

Houser, P. R., Shuttleworth, W. J., Famiglietti, J. S., Gupta, H. V., Syed, K. H., & Goodrich, D. C. (1998). Integration of soil moisture remote sensing and hydrologic modeling using data assimilation. *Water Resources Research*, *34*(12), 3405–3420.

Khaki, M., Ait-El-Fquih, B., Hoteit, I., Forootan, E., Awange, J., & Kuhn, M. (2017). A two-update ensemble Kalman filter for land hydrological data assimilation with an uncertain constraint. *Journal of Hydrology*, *555*, 447–462.

Khaki, M., Hoteit, I., Kuhn, M., Awange, J., Forootan, E., van Dijk, A. I. J. M., et al. (2017). Assessing sequential data assimilation techniques for integrating grace data into a hydrological model. *Advances in Water Resources*, *107*, 301–316.

Khaki, M., Schumacher, M., Forootan, E., Kuhn, M., Awange, J. L., & van Dijk, A. I. J. M. (2017). Accounting for spatial correlation errors in the assimilation of grace into hydrological models through localization. *Advances in Water Resources*, *108*, 99–112.

Kumar, S. V., Zaitchik, B. F., Peters-Lidard, C. D., Rodell, M., Reichle, R., Li, B., et al. (2016). Assimilation of gridded GRACE terrestrial water storage estimates in the North American land data assimilation system. *Journal of Hydrometeorology*, *17*, 1951–1972.

Lee, H., Seo, D. J., & Koren, V. (2011). Assimilation of streamflow and in situ soil moisture data into operational distributed hydrologic models: Effects of uncertainties in the data and initial model soil moisture states. *Advances in Water Resources*, *34*, 1597–1615.

Li, B., & Rodell, M. (2015). Evaluation of a model-based groundwater drought indicator in the conterminous us. *Journal of Hydrology*, *526*, 78–88.

Li, B., Rodell, M., Zaitchik, B. F., Reichle, R. H., Koster, R. D., & van Dam, T. M. (2012). Assimilation of GRACE terrestrial water storage into a land surface model: Evaluation and potential value for drought monitoring in western and central Europe. *Journal of Hydrology*, *446*, 103–115.

Lymburner, L., & Australia, G. (2011). of Agricultural, A.B., Economics, R.,. The national dynamic land cover dataset. Symonston, A.C.T. : Geoscience Australia. URL: geocat no.: 71069.

McVicar, T. (2011). Near-Surface Wind Speed. v8. Technical Report. CSIRO. Data Collec tion. https://doi.org/10.4225/08/589422252fa06

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*, 88–102.

Reichle, R. H., Crow, W. T., & Keppenne, C. L. (2008). An adaptive ensemble Kalman filter for soil moisture data assimilation. *Water resources research*, *44*, W03423. https://doi.org/10.1029/2007WR006357

Reichle, R. H., Koster, R. D., Dong, J., & Berg, A. A. (2004). Global soil moisture from satellite observations, land surface models, and ground data: Implications for data assimilation. *Journal of Hydrometeorology*, *5*, 430–442.

Reichle, R. H., Walker, J. P., Koster, R. D., & Houser, P. R. (2002). Extended versus ensemble Kalman filtering for land data assimilation. *Journal of Hydrometeorology*, *3*, 728–740. https://doi.org/10.1175/1525-7541(2002)003<0728:evekff>2.0.co;2

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. https://doi.org/10.1016/j.jhydrol.2014.08.008

Robock, A., Schlosser, C. A., Vinnikov, K. Y., Speranskaya, N. A., Entin, J. K., & Qiu, S. (1998). Evaluation of the AMIP soil moisture simulations. *Global and Planetary Change*, *19*, 181–208.

Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C. J., et al. (2004). The global land data assimilation system. *Bulletin of the American Meteorological Society*, *85*, 381–394. https://doi.org/10.1175/BAMS-85-3-381

Rodell, M., Velicogna, I., & Famiglietti, J. S. (2009). Satellite-based estimates of groundwater depletion in India. *Nature*, *460*, 999–1002. https://doi.org/10.1038/nature08238

Rowlands, D. D., Luthcke, S. B., Klosko, S. M., Lemoine, F. G. R., Chinn, D. S., McCarthy, J. J., et al. (2005). Resolving mass flux at high spatial and temporal resolution using GRACE intersatellite measurements. *Geophysical Research Letters*, *32*, L04310. https://doi.org/10.1029/2004GL021908

Schumacher, M., Kusche, J., & Döll, P. (2016). A systematic impact assessment of GRACE error correlation on data assimilation in hydrological models. *Journal of Geodesy*, *90*, 537–559.

Shokri, A., Walker, J., van Dijk, A. I. J. M., Wright, A., & Pauwels, V. (2017). Application of the patient rule induction method to hydrologic model parameter estimation and uncertainty quantification. *Hydrological Processes*, *32*, 1005–1025.

Slater, A. G., & Clark, M. P. (2006). Snow data assimilation via an ensemble Kalman filter. *Journal of Hydrometeorology*, *7*, 478–493.

Smith, B. S. (2013). Soil moisture monitoring with ground-based gravity data (PhD thesis), Monash University.

Strassberg, G., Scanlon, B. R., & Chambers, D. (2009). Evaluation of groundwater storage monitoring with the GRACE satellite: Case study of the High Plains aquifer, central United States. *Water Resources Research*, *45*, W05410. https://doi.org/10.1029/2008WR006892

Su, H., Yang, Z. L., Dickinson, R. E., Wilson, C. R., & 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. https://doi.org/10.1029/2009JD013035

Sun, C., Walker, J. P., & Houser, P. R. (2004). A methodology for snow data assimilation in a land surface model. *Journal of Geophysical Research*, *109*, D08108. https://doi.org/10.1029/2003JD003765

Swenson, S., & Wahr, J. (2006). Post-processing removal of correlated errors in GRACE data. *Geophysical Research Letters*, *33*, L08402. https://doi.org/10.1029/2005GL025285

Syed, T. H., Famiglietti, J. S., & Chambers, D. P. (2009). GRACE-based estimates of terrestrial freshwater discharge from basin to continental scales. *Journal of Hydrometeorology*, *10*, 22–40. https://doi.org/10.1175/2008JHM993.1

Tang, Q., Gao, H., Yeh, P., Oki, T., Su, F., & Lettenmaier, D. P. (2010). Dynamics of terrestrial water storage change from satellite and surface observations and modeling. *Journal of Hydrometeorology*, *11*, 156–170. https://doi.org/10.1175/2009JHM1152.1

Tangdamrongsub, N., Steele-Dunne, S. C., Gunter, B. C., Ditmar, P. G., & 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. https://doi.org/10.5194/hess-19-2079-2015

Tapley, B. D., Bettadpur, S., Ries, J. C., Thompson, P. F., & Watkins, M. M. (2004). GRACE measurements of mass variability in the Earth system. *Science*, *305*, 503–505.

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*, 1820–1840. https://doi.org/10.1002/2016WR019641

van Dijk, A. I. J. M. (2010). The Australian water resources assessment system (Technical Report 3. Landscape Model (version 0.5) Technical Description Technical Report): CSIRO: Water for a Healthy Country National Research Flagship.

van Dijk, A. I. J. M., Peña-Arancibia, J. L., & Bruijnzeel, L. A. (2012). Land cover and water yield: Inference problems when comparing catchments with mixed land cover. *Hydrology and Earth System Sciences*, *16*, 3461–3473.

van Dijk, A. I. J. M., Renzullo, L. J., & 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 Resources Research*, *47*, W11524. https://doi.org/10.1029/2011WR010714

van Dijk, A. I. J. M., Renzullo, L. J., Wada, Y., & Tregoning, P. (2014). A global water cycle reanalysis (2003-2012) merging satellite gravimetry and altimetry observations with a hydrological multi-model ensemble. *Hydrology and Earth System Sciences*, *18*, 2955–2973. https://doi.org/10.5194/hess-18-2955-2014

Vörösmarty, C. J., McIntyre, P. B., Gessner, M. O., Dudgeon, D., Prusevich, A., Green, P., et al. (2010). Global threats to human water security and river biodiversity. *Nature*, *467*, 555–561. https://doi.org/10.1038/nature09440

Wang, X., de Linage, C., Famiglietti, J., & Zender, C. S. (2011). Gravity recovery and climate experiment (GRACE) detection of water storage changes in the three gorges reservoir of China and comparison with in situ measurements. *Water Resources Research*, *47*, W12502. https://doi.org/10.1029/2011WR010534

Yeh, P. J. F., Swenson, S. C., Famiglietti, J. S., & Rodell, M. (2006). Remote sensing of groundwater storage changes in illinois using the Gravity Recovery and Climate Experiment (GRACE). *Water Resources Research*, *42*, W12203. https://doi.org/10.1029/2006WR005374

Zaitchik, B. F., Rodell, M., & 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. https://doi.org/10.1175/2007JHM951.1