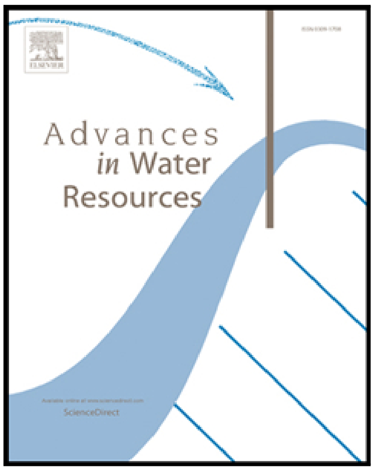

Assessing sequential data assimilation techniques for integrating GRACE data into a hydrological model

M. Khaki, I. Hoteit, M. Kuhn, J. Awange, E. Forootan, A. van Dijk, M. Schumacher, C. Pattiaratchi

PII: S0309-1708(16)30756-4
DOI: 10.1016/j.advwatres.2017.07.001
Reference: ADWR 2886

To appear in: *Advances in Water Resources*

Received date: 7 December 2016
Revised date: 2 July 2017
Accepted date: 3 July 2017

Please cite this article as: M. Khaki, I. Hoteit, M. Kuhn, J. Awange, E. Forootan, A. van Dijk, M. Schumacher, C. Pattiaratchi, Assessing sequential data assimilation techniques for integrating GRACE data into a hydrological model, *Advances in Water Resources* (2017), doi: 10.1016/j.advwatres.2017.07.001

Highlights

- We use GRACE data to improve a hydrological model estimations
- Data assimilation is used to ingrate observation into a model
- We apply stochastic and deterministic ensemble-based Kalman filters (EnKF) and Particle filter
- Filters performances are compared to reach the best result
- Independent in-situ measurements are used to evaluate the results

Assessing sequential data assimilation techniques for integrating GRACE data into a hydrological model

M. Khaki[a,1], I. Hoteit[b], M. Kuhn[a], J. Awange[a], E. Forootan[a,c], A. van Dijk[d], M. Schumacher[e], C. Pattiaratchi[f]

[a] *Western Australian Centre for Geodesy and The Institute for Geoscience Research, Curtin University, Perth, Australia.*

[b] *King Abdullah University of Science and Technology, Thuwal, Saudi Arabia.*

[c] *School of Earth and Ocean Sciences, Cardiff University, Cardiff, UK.*

[d] *Fenner School of Environment and Society, the Australian National University, Canberra, Australia.*

[e] *Institute of Geodesy and Geoinformation, University of Bonn, Nussallee 17, 53115 Bonn, Germany.*

[f] *School of Civil, Environmental, and Mining Engineering / UWA Oceans Institute, The University of Western Australia, Crawley, Australia.*

Abstract

The time-variable terrestrial water storage (TWS) products from the Gravity Recovery And Climate Experiment (GRACE) have been increasingly used in recent years to improve the simulation of hydrological models by applying data assimilation techniques. In this study, for the first time, we assess the performance of the most popular data assimilation sequential techniques for integrating GRACE TWS into the World-Wide Water Resources Assessment (W3RA) model. We implement and test stochastic and deterministic ensemble-based Kalman filters (EnKF), as well as Particle filters (PF) using two different resampling approaches of Multinomial Resampling and Systematic Resampling. These choices provide various opportunities for weighting observations and model simulations during the assimilation and also accounting for error distributions. Particularly, the deterministic EnKF is tested to avoid perturbing observations before assimilation (that is the case in an ordinary EnKF). Gaussian-based random updates in the EnKF approaches likely do not fully represent the statistical properties of the model simulations and TWS observations. Therefore, the fully non-Gaussian PF is also applied to estimate more realistic updates. Monthly GRACE TWS are assimilated into W3RA covering the entire Australia. To evaluate the filters performances and analyze their impact on model simulations, their estimates are validated by independent in-situ measurements. Our results indicate that all implemented filters improve the estimation of water storage simulations of W3RA. The best results are obtained using two versions of deterministic EnKF, i.e. the Square Root Analy-

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

[1]Contact details: Western Australian Centre for Geodesy and The Institute for Geoscience Research, Curtin University, Perth, Australia, Email: Mehdi.Khaki@postgrad.curtin.edu.au, Tel: 0061410620379

sis (SQRA) scheme and the Ensemble Square Root Filter (EnSRF), respectively improving the model groundwater estimations errors by 34% and 31% compared to a model run without assimilation. Applying the PF along with Systematic Resampling successfully decreases the model estimation error by 23%.

Keywords: Data assimilation, GRACE, Hydrological modelling, Kalman filtering, Particle filtering.

1. Introduction

Hydrological models offer important tools for simulating and predicting hydrological processes at global (e.g., Doll et al., 2003; Hunt, 2006; Coumou and Rahmstorf, 2012; van Dijk et al., 2013) and regional (e.g., Chiew et al., 1993; Wooldridge and Kalma, 2001; Christiansen et al., 2007; Huang et al., 2016) scales. Models are still being developed to simulate all available hydrological processes (e.g., groundwater recharge) and the inclusion of all interactions between water cycle components (e.g., evapotranspiration, precipitation, and runoff). Currently, the most important deficiencies in hydrological models are caused by a high level of uncertainties in imperfect modelling of complex water cycle processes, data deficiencies on both temporal and spatial resolutions (e.g., limited ground-based observations), uncertainties in input and forcing data, and uncertainties of (unknown) empirical model parameters (Vrugt et al., 2013; van Dijk et al., 2011, 2014). Since making models more complex introduces ever increasing model parameters that cannot be well interpreted and makes computations more expensive, a logical step to address these limitations is the assimilation of observations into models (e.g., McLaughlin, 2002; Zaitchik et al., 2008; van Dijk et al., 2014). Data assimilation techniques have found increasing interests with the availability of new data sources, such as those derived from satellite remote sensing observations. For example, time-variable gravity fields from the Gravity Recovery And Climate Experiment (GRACE) mission (Tapley et al., 2004) can be converted to terrestrial water storage (TWS) fields, a fundamental parameter of the water cycle that might be used to reduce uncertainties in hydrological models.

Data assimilation is a procedure that constrains the dynamic of a model with available observations in order to improve its estimates (Bertino et al., 2003). The solution of the data assimilation problem is based on the Bayes' rule (Jazwinski, 1970; van Leeuwen and Evensen, 1996), which basically computes the Probability Density Function (PDF) of the state, i.e., the model variable of the system that should be estimated, given the data. The updated distribution is then propagated with the model to the time of the next available observation to obtain the prior PDF. In the case of a nonlinear or non-Gaussian system (as it is the case for hydrological models), it is not possible to analytically derive the posterior (analysis) PDF of the state (Hoteit et al., 2008; Vrugt et al., 2013). The Bayesian estimation problem, therefore, needs to be solved numerically, using either variational smoothing or sequential filtering methods (Subramanian et al., 2012).

Variational methods look for the model trajectory that best fits the data by minimizing a chosen cost function that measures the misfit between the model state and the observations (Talagrand and Courtier, 1987). These methods require coding and executing an adjoint model, which is very demanding in terms of human and computational resources (Hoteit et al., 2005). Furthermore, variational methods do not provide an efficient framework for updating the estimating statistics during the data assimilation process (Courtier et al., 1994; Kalnay, 2003). In contrast, sequential techniques process the data as they become available following two steps including a forecast step to propagate the distribution forward in time and an analysis step to update the distribution with the newly available observation. Monte Carlo methods are commonly used in the forecast step (based on ensembles or particles) and Kalman (Ensemble Kalman filtering) or point-mass weight (Particle filtering) updates are applied in the analysis step (Evensen, 2009; Hoteit et al., 2012). Sequential methods do not require an adjoint and are becoming increasingly popular because of their reasonable computational requirements (Hoteit et al., 2002; Bertino et al., 2003; Robert et al., 2006).

The Particle filter (PF) is based on a point-mass (particle) representation of the system state's PDF. It forecasts the PDF by propagating the particles forward in time. At the analysis time, the state PDF is updated by assigning new weights to the particles based on incoming observations (Doucet et al., 2001; Pham, 2001; Hoteit et al., 2012). The fundamental problem of this technique is the degeneracy phenomenon of its particles, with only very few particles carrying most of the weights (Subramanian et al., 2012). Moreover, errors in the assimilated observations may propagate to the estimated distribution because the method was not designed to improve the structure of the model (Hoteit et al., 2008; Smith et al., 2008). This problem is addressed by the Ensemble Kalman filters (EnKFs), which assume a Gaussian forecast PDF at the analysis time, so a Kalman update-step is applied to the particles (Hoteit et al., 2015). This allows an efficient implementation of the Bayesian filtering approach for data assimilation into large systems using small ensembles (van Leeuwen and Evensen, 1996; Hoteit et al., 2008).

EnKFs can be classified into stochastic and deterministic filters, depending on whether the observations are perturbed with noise before assimilation, or not (Tippett et al., 2003; Hoteit et al., 2015). In the stochastic EnKF, each ensemble member is updated with perturbed observations, readily providing an analysis ensemble for the next filtering cycle. In contrast, a deterministic EnKF updates only the mean and the covariance of the ensembles exactly as in the Kalman Filter, and thus require a resampling step to generate a new analysis ensemble. The resampling step is not unique, and as such several deterministic EnKFs have been proposed (Sun et al., 2009; Hoteit et al., 2015).

Sequential filtering methods have been extensively applied and compared in oceanic and atmospheric applications (Garner et al., 1999; Elbern and Schmidt, 2001; Bennett, 2002; Kalnay, 2003; Schunk et al., 2004; Lahoz, 2007; Zhang et al., 2012; Altaf et al., 2014). In hydrological studies, data assimilation has been used to estimate different water compartments, such as soil moisture (e.g., Reichle et al., 2002; Brocca et al., 2010; Renzullo et al., 2014) and surface water storage (e.g., Alsdorf et al., 2007; Neal et al., 2009; Giustarini et al., 2011). However, the efficiency of various filtering methods in dealing with remotely sensed data in hydrology has not been fully investigated (McLaughlin, 2002; Schumacher et al., 2016).

Global terrestrial water storage data derived from the GRACE satellite mission can be now employed to improve the behaviour of hydrological models (e.g., Zaitchik et al., 2008; Tangdamrongsub et al., 2015; Thomas et al., 2014; van Dijk et al., 2014; Eicker et al., 2014; Reager et al., 2015), providing unprecedented temporal and spatial coverage. For instance, Zaitchik et al. (2008) demonstrated the relevance of GRACE data in improving the estimation of groundwater variability over the four major sub-basins of the Mississippi through data assimilation into the Catchment Land Surface Model using an ensemble Kalman smoother. Houborg et al. (2012) investigated drought conditions in North America through GRACE data assimilation. The developed GRACE-based drought indicators in the USA led to an improved monitoring of soil moisture and groundwater conditions of deep layers. The impact of GRACE error correlation structure on the assimilation of GRACE data was very recently studied by Schumacher et al. (2016). Yet, to the best of our knowledge, however, a comparison with the application of different sequential filtering methods for assimilating GRACE TWS in models has not been fully explored.

In this study, we investigate the performance of the most common sequential filtering techniques for data assimilation using the hydrological model of the World-Wide Water Resources Assessment (W3RA; van Dijk, 2010) over Australia. The amount of rainfall in Australia, especially over its northern and eastern parts, is low in comparison to other inhabited continents on Earth leading to prolonged drought in the interior regions (Forootan et al., 2016). Hence, accurate estimation of water storages (e.g., using hydrological models) is necessary to manage water resources in this region. Here, different filters are used to assimilate GRACE TWS into W3RA to improve its estimates. Both stochastic and deterministic EnKFs are tested and their performances are compared against two standard Particle filters. We applied the standard EnKF and its deterministic variants, including, the Square Root Analysis (SQRA) scheme following Evensen (2004) and Schumacher et al. (2016), the Ensemble Transform Kalman Filter (ETKF, Bishop et al., 2001), the Deterministic EnKF (DEnKF, Sakov and Oke, 2008), and the Ensemble Square-Root Filter (EnSRF, Whitaker and Hamill, 2002). We also implement the static-ensemble variant of the EnKF, the Ensemble Optimal Interpolation (EnOI, Evensen, 2003), in an attempt to reduce the computational burden. To mitigate the deficiency that may arise from limited ensemble sizes and knowledge of model errors' statistics (Anderson et al., 2007; Oke et al., 2007), covariance inflation (e.g., Anderson et al., 1999, 2007; Ott et al., 2004) and localization techniques (e.g., Bergemann and Reich, 2010; Hamill and Snyder, 2002) are applied. The performance of these ensemble filters is assessed against two nonlinear Particle filters based on two different resampling strategies: (i) Multinomial Resampling and (ii) Systematic Resampling techniques (Arulampalam et al., 2002). The summary of applied filters in this study is presented in Table 1. The results of assimilations are evaluated by comparing their estimates against independent groundwater in-situ measurements over the Murray-Darling basin and measurements from the moisture-monitoring network in the Murrumbidgee catchment in New South Wales, Australia.

TABLE 1

2. Model and Datasets

2.1. W3RA

The World-Wide Water Resources Assessment (W3RA), based on the Australian Water Resources Assessment system (AWRA) model (version 0.5) is used in this study (http://www.wenfo.org/wald/data-software/). The model was first developed in 2008 by the Commonwealth Scientific and Industrial Research Organisation (CSIRO) to monitor, represent and forecast Australian terrestrial water cycles. The W3RA is a grid-distributed biophysical model that simulates landscape water stores in the vegetation and soil systems (van Dijk, 2010). The $1° \times 1°$ global daily fields of minimum and maximum temperature, downwelling short-wave radiation, and precipitation from Princeton University (http://hydrology.princeton.edu) are used for meteorological forcing data (Sheffield et al., 2006). This one-dimensional grid-based water balance model represents the water balance of the soil, groundwater and surface water stores in which each cell is modelled independently of its neighbours (van Dijk, 2010; Renzullo et al., 2014). The model state is composed of the $1° \times 1°$ W3RA model storages of the top, shallow root and deep root soil layers, groundwater storage, and surface water storage in a one-dimensional system (vertical variability). In this study, we use W3RA providing daily model states for the period of February 2002 to December 2012. More detailed information on the W3RA model can be found in van Dijk (2010).

2.2. GRACE-derived Terrestrial Water Storage

Here, we use monthly GRACE level 2 (L2) products along with their full error information between February 2002 to December 2012 as provided by the ITSG-Grace2014 gravity field model (Mayer-Gurr et al., 2014). The GRACE monthly Stokes' coefficients are truncated at spherical harmonic degree and order 90, which resulting in approximately ∼300 by 300 km spatial resolution at the equator.

Following Swenson et al. (2008), degree 1 coefficients are replaced to account for movements of the Earth's centre of mass (i.e., realized by a set of tracking stations on the surface of the Earth). Degree 2 and order 0 (C20) coefficients from GRACE are not well determined (e.g., Tapley et al., 2004; Tregoning et al., 2012) and are replaced by more reliable estimations of the Satellite Laser Ranging solutions (Cheng and Tapley, 2004). Correlated noise exists in L2 products due to anisotropic spatial sampling, instrumental noise (K-band ranging system and

GPS), and temporal aliasing caused by the incomplete reduction of short-term mass variations (Forootan et al., 2014). These errors are reduced by smoothing based on a Gaussian averaging kernel with 300 km half radius and destriping following Swenson and Wahr (2006). However, the smoothing may cause signal attenuation (Klees et al., 2008) and can result in considerable spatial leakage, such as the apparent movement of masses from one region to another (Chen et al., 2007) especially over coastlines (see examples within Australia in e.g., Brown and Tregoning, 2010; Forootan et al., 2012). In order to address this issue, following Swenson and Wahr (2002), we apply an isotropic kernel using a Lagrange multiplier filter to best balance signal and leakage errors over the basin of interest.

An additional post-processing step is applied to convert the filtered L2 gravity fields (after removing the mean field of study period) to gridded TWS fields ($1^{\circ}\times1^{\circ}$) following Wahr et al. (1998). The GRACE TWS data are gridded at the same spatial $1^{\circ}\times1^{\circ}$ resolution of W3RA resulting in 794 grid points for Australia that covers an area of 7.692 million km^2 located between $10^{\circ}S$ and $46^{\circ}S$ latitude, and $110^{\circ}E$ and $160^{\circ}E$ longitude. GRACE data provide changes in TWS while W3RA produces absolute TWS. Accordingly, the mean TWS for the study period is taken from W3RA and is added to the GRACE TWS change time series in order to obtain absolute values in accordance with the model (Zaitchik et al., 2008). In addition, the monthly full error information of the Stokes' coefficients is used to construct an observation error covariance matrix for the GRACE TWS fields (Eicker et al., 2014; Schumacher et al., 2016).

2.3. In-situ data

For validating the assimilation results, we use in-situ groundwater level data that are collected over the Murray-Darling Basin. The independent in-situ measurements from the model and observations are provided by New South Wales Government (NSW) groundwater archive (http://waterinfo.nsw.gov.au/pinneena/gw.shtml). Monthly well measurements are acquired and time series of groundwater storage anomalies are generated. Measurements with data gaps and those without showing seasonal variations are flagged (we assume these belong to confined aquifers) and are thus excluded (Houborg et al., 2012; Tangdamrongsub et al., 2015). Selected bore-water levels are then converted to variations in groundwater (GW) storage. To this end, instead of using specific yield estimates (Rodell et al., 2007; Zaitchik et al., 2008) that is not available in the region, TWS variation from GRACE and GLDAS soil moisture are used to scale the observed head following Tangdamrongsub et al. (2015). Tregoning et al. (2012) show that this approach can be used to find a scaling factor over the Canning Basin and Murray Basin in Australia. The scaled in-situ groundwater level fluctuations are then used to assess the assimilation results.

In addition, in-situ measurements of the moisture-monitoring network (http://www.oznet.org.au/) in Murrumbidgee catchment (Smith et al., 2012) are used to evaluate the results. These data are known as the OzNet network, which provides long-term records of measuring volumetric soil moisture at various soil depths at 57 locations across the Murrumbidgee catchment area. Following Renzullo et al. (2014), we averaged the measurements into a daily scale and use 0–8 cm to evaluate the estimated model top-layer soil moisture and the 0-30 cm and 0-90 cm measurements for the evaluation of the model shallow root-zone soil moisture estimation. The distribution of the in-situ moisture network, as well as in-situ groundwater stations, are shown in Figure 1.

FIGURE 1

3. Filtering Methods and Implementation

The Bayesian filtering procedures are selected here for data assimilation (Jazwinski, 1970; van Leeuwen and Evensen, 1996). The analytical process conditions a prior PDF of the state with available observations to compute the posterior PDF based on Bayes' rule (Koch, 2007) in two steps; (1) forecasting the state PDF using a dynamical model and (2) updating the forecast PDF by assimilating observations using Bayes' rule. In the case of a linear system with Gaussian noise, the popular KF provides the Bayesian filtering solution by computing the first two moments of the state PDF, which remains always Gaussian (van Leeuwen and Evensen, 1996). This two-step process is repeated whenever new observations become available. The basic KF equations are given by Kalman (1960) starting from an analysis of the state, X_t^a, and the associated error covariance, P_t^a, at a given time t. These can be summarized as:

1) The forecast step consists of the evolution of the state estimate and its error covariance matrix with a linear dynamical model (M) to the assimilation step (the time of the next available observation),

$$X^f_{t+\Delta t} = MX^a_t + \eta, \tag{1}$$

$$P^f_{t+\Delta t} = MP^a_t M^T + Q, \tag{2}$$

where $X^f_{t+\Delta t}$ refers to the forecast state (X^f) at time $t + \Delta t$, with Δt represents the model time step, and T is the transpose index. In Eq.1, η is the process noise, which is drawn from $N(0, Q)$ with covariance matrix Q, and $P^f_{t+\Delta t}$ (in Eq.2) denotes the forecast error covariance (P^f) at time $t + \Delta t$.

2) The analysis step updates the forecast state using new incoming observations Y that are related to the state vector by the linear observation operator (H) as,

$$Y = HX + \epsilon, \tag{3}$$

where ϵ is the measurement noise. The analysis state (X^a) is then computed using

$$X^a = X^f + K(Y - HX^f), \tag{4}$$

$$P^a = (I - KH)P^f, \tag{5}$$

$$K = P^f H^T (HP^f H^T + R)^{-1}, \tag{6}$$

where K refers to the Kalman gain, R is the observation error covariance matrix, and I denotes the identity matrix.

The KF algorithm is not suited for high-dimensional or non-linear systems (Pham, 2001). The ensemble Kalman filter provides an efficient alternative for the implementation of the KF with these systems by representing the first two-moments of the state using a sample of state vectors, called ensemble. The forecast state and covariance matrix in Eq.1 and Eq.2 are then estimated as the sample mean and covariance of the ensembles members $X^i, i = 1 \ldots N$:

$$\bar{X}^f = \frac{1}{N}\sum_{i=1}^{N} X^{f,i}, \tag{7}$$

$$P^f = \frac{1}{N-1}\sum_{i=1}^{N}(X^{f,i} - \bar{X}^f)(X^{f,i} - \bar{X}^f)^T = \frac{1}{N-1}A^f A^{fT}. \tag{8}$$

$\bar{X}^f$ is the forecast ensemble mean and A^f ($A^f = [A^{f,1} \ldots A^{f,N}]$) is the forecast ensemble of anomalies (perturbations; $A^{f,i} = X^{f,i} - \bar{X}$). When a new observation is available, the forecast ensemble is then updated with the data using Eq.4, as in the KF. Several ensemble Kalman filters have been proposed, all sharing the same forecast step in which an available analysis ensemble ($[X^{a,1} \ldots X^{a,N}]$) is propagated forward with the model. The analysis step based on the KF, however, can be applied stochastically or deterministically.

3.1. Stochastic Ensemble Kalman Filter (EnKF)

The analysis step of the Stochastic EnKF updates each ensemble member with a perturbed observation written as,

$$X^{a,i} = X^{f,i} + K(Y^i - HX^{f,i}), \quad i = 1 \ldots N, \tag{9}$$

where $Y^i = Y + \varepsilon^i$, with ε^i a random error sampled from $N(0, R)$. The use of perturbed observations in the EnKF results in an analysis error covariance that matches that of the KF, in a statistical sense (Hoteit et al., 2012). The advantage of the stochastic update is that it readily provides a randomly sampled ensemble from the Gaussian-assumed state analysis PDF for the next forecast cycle (Hoteit et al., 2015). However, as illustrated by Whitaker and Hamill (2002), sampling error can be reflected in the EnKF background covariance matrix, especially for the small-size ensembles. This could be particularly pronounced when a large number of (independent) observations are assimilated (Nerger, 2004), as the observation covariance cannot be properly sampled with a small ensemble (Hoteit et al., 2015).

3.2. Deterministic Ensemble Kalman Filters

Instead of updating each forecast member separately, deterministic EnKFs (DEnKFs) update the forecast ensemble in two steps, first the ensemble-mean and then the ensemble perturbations (Tippett et al., 2003) are calculated so that the sample mean and covariance of the updated ensemble exactly match those of the Kalman filter in Eq.4 and Eq.5.

Various methods have been proposed in order to update the ensemble perturbations. SQRA resamples the new ensemble perturbations (A^a) from the forecast ensemble perturbations (A^f) as,

$$A^a = A^f V \sqrt{I - \Sigma^T \Sigma \Theta^T}, \tag{10}$$

where Σ is computed by applying the following singular value decomposition (SVD),

$$U\Sigma V^T = SVD(\Lambda^{-\frac{1}{2}} Z^T HA), \tag{11}$$

$$Z\Lambda Z^T = (HP^f(H)^T + R)^{-1}, \tag{12}$$

where Θ in Eq.10 being a random orthogonal matrix for redistribution of the variance among the ensemble members (see Evensen, 2004, 2007, for more details). This is very similar to the random rotation that has been introduced in the context of the Singular Evolutive Extended Kalman (SEEK) filter (Pham, 2001; Hoteit et al., 2002).

ETKF introduces a transformation matrix to directly compute the analysis ensemble perturbations from their forecast counterparts,

$$A^a = A^f.T, \tag{13}$$

where $T = U(I+\Lambda)^{-1/2}$, with U and Λ, respectively being the orthogonal and diagonal matrices computed from an eigenvalue decomposition of $(HX^f)^T R^{-1}(HX^f)$.

DEnKF and EnSRF adopt a similar analysis step to the EnKF in the sense that they compute the analysis perturbations from the forecast perturbations by updating each ensemble perturbation with a Kalman-like update step. To match the KF covariance matrix by an ensemble of perturbations, DEnKF computes a first-order approximation of the Kalman gain (Sakov and Oke, 2008). This approximate gain $\tilde{K}$ is then used to compute the analysis perturbations as,

$$A^a = A^f - \frac{1}{2} KHA^f. \tag{14}$$

EnSRF exploits the serial formulation of the KF analysis step in which the observations are assimilated each at a time to compute the analysis perturbations that exactly match the KF covariance using the modified gain (αK) with,

$$\alpha = \left(1 + \sqrt{\frac{R}{HP^f H^T + R}}\right)^{-1}. \tag{15}$$

This requires the observations to be uncorrelated, which can always be satisfied by scaling the observations with the square-root inverse of the observational error covariance matrix (Hoteit et al., 2015).

Another form of ensemble Kalman filtering is the so-called Ensemble Optimal Interpolation (EnOI) scheme, which is basically the EnKF, but without an update of the ensemble anomalies. More precisely, EnOI only updates the forecast state with a Kalman gain computed from a preselected static ensemble. The main advantage of not updating the ensemble is of course to reduce the computational load, but it can also be beneficial to retain the spread of the ensemble and to enforce climatological smoothness in the update step. EnOI can be stochastic or deterministic (Hoteit et al., 2002). Here we only test the more standard stochastic variant (Evensen, 2003).

3.3. Particle Filtering

Particle filtering is also a sequential Monte Carlo method that was originally proposed by Gordon et al. (1993) and has since been applied in numerous studies (Doucet, 1998; Arulampalam et al., 2002). The idea is to represent the state PDF by a set of weighted particles (Arulampalam et al., 2002), hence the name Particle Filter (Gordon et al., 1993; Doucet, 1998), which is similar to the ensemble members in the EnKF but with non-uniform weights. The state PDF is then decomposed as,

$$P(X_t|Y_{1:t}) \approx \sum_{i=1}^{N} \omega_t^i \delta(X_t - X_t^i), \tag{16}$$

where $\{X_t^i; i = 1 \dots N\}$ are the particles at time t, observations between time 1 and t are denoted by $Y_{1:t}$, ω_t^i are the weights of the particles, and δ is the Dirac function. In the forecast step, the PF just integrates the particles forward with the model, exactly as the EnKF, and their weights remain the same. In the analysis step, only the weights, and not the particles, are updated with the incoming observation as,

$$\omega_t^i = \frac{P(y_t|X^i_{t|t-1})}{\sum_j P(y_t|X^j_{t|t-1})}. \tag{17}$$

The PF suffers from the degeneracy problem in which the weights of all particles become negligible except only for a very few, requiring a prohibitive number of particles to prevent particles collapse (Arulampalam et al., 2002). Degeneracy can be mitigated using the so-called resampling technique (Doucet et al., 2005), which resamples a new set of particles with uniform weights after every update step based on the analysis PDF. In this study, we consider two of the most common resampling techniques: the Particle filter with Multinomial Resampling (PFMR) and Particle filter with Systematic Resampling (PFSR), as proposed by Doucet et al. (2005). PFMR is the most straightforward resampling scheme, where N independent random numbers ($u \sim U(0,1)$) are generated to select a particle from the old set. PFSR, which is also called universal sampling, draws only one random number $u_1 \sim U(0,1/N)$ and the remaining $N-1$ numbers are then calculated from u_1 (Doucet et al., 2005) as,

$$U_i = u_1 + \frac{(i-1)}{N}, \quad i = 2 \dots N. \tag{18}$$

These are then used to select a new set of particles according to the multinomial distribution (Hol et al., 2006).

PF has been applied in few hydrological studies. Among them, Moradkhani et al. (2005) investigated the relevance of the PF for estimating the joint posterior distribution of the parameters and state. In another effort, Moradkhani et al. (2012) proposed a modified version of the PF, focusing on enhancing the sampling of the posterior with Markov chain Monte Carlo (MCMC) moves. Plaza et al. (2012) used the Sequential Importance Sampling with Resampling (SISR) Particle filter for soil moisture assimilation and focused on the consequent effect on baseflow generation. Existing studies focused on the different implementations of PF using various resampling techniques. However, a comparison between PFs performances with diverse resampling techniques and EnKFs has not been investigated yet in hydrology. Figure 2 shows a summary of the steps and filters applied for data assimilation in this study.

FIGURE 2

3.4. Filter Implementation

An experimental framework is developed in order to assess the relevance and efficiency of the filtering techniques presented in the previous section for assimilating GRACE data into the W3RA model. All filters are implemented under identical conditions, using the same spatial scales (1°×1°) for both the W3RA and the GRACE TWS, and daily temporal scales for the W3RA and monthly for the GRACE data. W3RA is integrated to simulate water storages over Australia using monthly sequential assimilation cycles of GRACE data applied at the middle of each month.

Several steps need to be undertaken before assimilating GRACE-derived TWS into the

W3RA model. Initial ensemble members (particles) are first generated by perturbing the three most important forcing variables including precipitation, temperature, and radiation using their reported error characteristics (Sheffield et al., 2006). Monte Carlo sampling of multivariate normal distributions with the errors representing the standard deviations of the forcing sets are used to produce an ensemble (see details in Renzullo et al., 2014). Different ensemble sizes (30-120) and their spread are tested which the ensemble with 72 members (72 to 120 are suggested by Oke et al., 2008) shows promising performance and is used in this study. The model is integrated forward for two years (January 2001 to January 2003) using perturbed meteorological forcing datasets and provided a set of 72 different states at the beginning of 2003 (study period), considered as the initial ensemble (with 72 members). The same initial ensemble is used for all the filters.

We use two tuning techniques of ensemble inflation and localisation in order to enhance the assimilation performance of all EnKFs. Ensemble inflation uses a small coefficient (i.e., 1.12 in our study; Anderson et al., 2001) to inflate prior ensemble deviation from the ensemble-mean to increase their variations and alleviate the inbreeding problem (Anderson et al., 2007). Another auxiliary technique that has been proved to be helpful when using limited ensemble size is localisation, initially proposed by Houtekamer and Mitchell (2001). We choose to use a Local Analysis (LA) scheme which works by restricting the impact of a given measurement in the update step to the points located within a certain distance from the measurement location (Evensen, 2003; Ott et al., 2004). Different localization lengths are applied to reach the best case (i.e., 5°). In terms of computational cost, all implemented filters are required more or less the same CPU (central processing unit) time (when implemented with the same number of members/particles), with the forecast step of the ensemble being the most demanding.

4. Results

In this section, we review and analyze the performance of all the selected filtering techniques based on various factors. The implemented filters include (stochastic) EnKF, ETKF, SQRA, DEnKF, EnSRF, EnOI and PF with Multinomial (PFMR) and Systematic (PFSR) resampling. In addition to improving the estimation of the system state and quantifying the associated uncertainties, a suitable data assimilation technique is expected to keep the model system stable during the assimilation process after incorporating GRACE data. These are provided at coarse temporal and spatial scales in comparison to the W3RA model, leading to only one assimilation step every 30 model time-steps and providing information at about three times less than the model grid resolution. Our analysis is organized into two parts; we first examine the filters performance by comparing their estimates (analysis and forecasts) against the assimilated GRACE data over the whole study area as well as the independent in-situ measurements over the Murray-Darling River Basin as well as Murrumbidgee catchment. We also compare the filters estimates with the outputs of a model-free run (open-loop) that is integrated with the filters initial condition without assimilation to evaluate the impact of assimilating GRACE data on the model behavior. Next, the filters behaviors in terms of ensemble spread and the impact of assimilation on the forecast and analysis error covariances are investigated.

4.1. Assessment with GRACE and in-situ data

Spatial correlation maps with high correlations may suggest that the applied filtering method efficiently incorporates GRACE data into the model (Figure 3). The correlation between the model TWS outputs without assimilation and the GRACE data range between 0.11 and 0.64, with the highest correlations in the northern region and the lowest in the southern region. All filters significantly improve the estimates correlations to the data after assimilation with some filters leading almost to the perfect correlation with the data (e.g., EnSRF). The model is not able to maintain this high correlation during the forecast and the 30-day assimilation window, with the correlations mainly decreasing in the center and southern regions. After monitoring the impact of observations on the model states throughout the study period, it is found that this effect is decreasing gradually (approximately 3-5 days to lose more than 10%) by comparing the correlation of the model states with and without assimilation. This mostly refers to the daily effects of the perturbed forcing sets on model estimations and may suggest that using the denser observation (temporally) could preserve assimilated information during the study. The level of improvement in correlations, however, is different for each filter. For instance, ETKF, SQRA, and PFSR lead to higher correlations with GRACE-derived TWS, suggesting that these methods better reflect the observations in the state estimates. Overall, EnKFs seem to perform better than PFs except only for DEnKF which shown no remarkable impact on the model behavior after assimilation of GRACE data.

FIGURE 3

Those methods with the highest correlations to GRACE data lead, as expected, also the least estimation errors (Figure 4). The largest errors are found in the northern and southern parts of the domain (Figure 4a), with some of the filters not able to improve remarkably the model behavior over these areas. TWS variations are generally higher in the northern part of the study area with larger amplitudes especially during monsoonal seasons (Awange et al., 2009; Seoane et al., 2013). The model seems unable to predict these amplitudes due to larger estimated errors even though it performs better in predicting their phases considering high correlations in this area. SQRA, EnSRF, and to some extent ETKF, significantly decrease the estimation error over the whole domain. This is very important because these filters are able to incorporate most of the GRACE signals into the model.

FIGURE 4

The Root-Mean-Squared Errors (RMSE) time series between the GRACE TWS and filters estimates are calculated (Figure 5). In all cases, the analysis step decreases the RMSE with respect to the forecast. Nevertheless, the RMSE resulting from DEnKF, EnOI, and PFMR are significantly larger, indicating that these methods are not able to improve the model behavior after incorporating GRACE data as the rest of the filters do. Estimates by all filters have the largest error in some periods (e.g., July and October), which may be caused by uncertainties in forcing sets. The RMSEs from SQRA and EnSRF are smaller in comparison to the rest of the filters. The smaller average errors during the study period prove the more stable performance of SQRA and EnSRF. Results in Figures 4 and 5 suggest that the deterministic SQRA, ETKF, EnSRF filters, and to less extent PFSR, are more efficient at assimilating GRACE data. This might be due to the fact that for the stochastic ensemble filters perturbations of the observations have to be generated that introduce an additional uncertainty source to the analysis step and might result in larger discrepancies to the assimilated observations compared to the deterministic filters. A summary of the filters' performance, including the coefficient of determination (R^2) and RMSE in comparison to the assimilated observations (shown in Table 2) indicates higher correlation (84% average) and smaller RMSEs (35% average improvement) in the analysis step for all the filters. The maximum improvements regarding the achieved RMSE are achieved by EnSRF and SQRA as 58.88% and 55.17% respectively.

FIGURE 5

TABLE 2

We investigate the performances of the filtering methods through comparison with the independent groundwater in-situ data over the Murray-Darling basin (cf. Section 2.3). We use the 54 in-situ measurements over the Murray-Darling basin for a grid comparison with the estimated GW (Figure 6). The filters estimates are spatially interpolated to the nearest observation bore. For each filter, the average RMSEs (over all 54 in-situ data) of the forecast state (red) and analysis state (blue) are determined. As for the assimilated GRACE data (cf. Figure 5), all the filters decrease the RMSE with respect to the in-situ data, with the largest errors resulting from DEnKF, EnOI, and PFMR. Furthermore, the average RMSEs are smaller in SQRA and EnSRF. The similar behavior of the filters in the analysis steps can be found in Figure 6 as in Figure 5. For some months (e.g., March and July), the larger errors can be seen in Figure 6 which are not existed in Figure 5. This can be associated to either an incompatibility between groundwater in-situ measurements and GRACE data or the absent water compartment terms such as the surface water storage in the model and in-situ data for the second assessment (Figure 6).

FIGURE 6

The relationship between the estimated states and both GRACE data and in-situ measurements (Figure 7) demonstrates the filters capability to dynamically propagate the information extracted from GRACE data into system variables. In agreement with the previous results, the best performances are obtained using SQRA, EnSRF, and ETKF (Figures 6 and 7).

FIGURE 7

The R^2 coefficient and RMSE results are summarized in Table 3 as another measure of the filter performances. For each filter, 54 error time series are computed (i.e., for each individual well), and their averages are then used to calculate R^2 and RMSE. The results of all the filters summarized in Table 3 show improvements (by 35% average) in the analysis steps in all cases, similar to the assessment against GRACE data (cf. Table 2). SQRA and EnSRF lead to the highest correlations to the in-situ measurements of R^2, i.e., 0.75 and 0.72 respectively. These filters also provide the best estimates in terms of estimation error, while DEnKF and to a lesser degree EnOI have the highest RMSEs. The PFs, on the other hand, especially using the systematic resampling technique exhibit a reasonably good performance. In terms of the assessment results against GRACE data, deterministic filters provide the best performance (except for DEnKF), generally better than the stochastic EnKF. Overall, SQRA and EnSRF seem to be the most efficient for assimilating GRACE data into W3RA.

TABLE 3

The correlations between model estimations and OzNet data also indicate the superiority of the successful methods in previous assessments (Table 4). Note that considering the difference between W3RA estimations (i.e., column water storage measured in mm) and the OzNet measurements (i.e., volumetric soil moisture) and the fact that converting the model output into volumetric units may introduce bias (Renzullo et al., 2014), only correlation analysis is assumed here. After estimating correlations for each individual layer, we determine an average correlation for the total soil column (cf. Table 4). The higher correlations are found in analysis steps with the average of 74% in comparison to forecast steps (59%). The highest correlation to the OzNet soil moisture measurements belongs to EnSRF with R^2 0.84. SQRA also demonstrate a significant impact on model estimations with the 35% correlation improvement. The weakest performance with R^2 0.48 and 0.57 in forecast and analysis step respectively, is achieved from DEnKF. These results prove the capability of EnSRF and SQRA in improving non-assimilated model states through data assimilation.

TABLE 4

4.2. Error Analysis

Analyzing the filters sampled error covariance, particularly the ensemble spread is important to understand the filters behaviors and performances. The performance of ensemble-based filters relies on their ability to represent and propagate the error statistics, which of course depends on how the ensemble members are sampled and updated at the analysis steps (Sun et al., 2009). We assess the evolution of the ensemble spread and the error covariance matrix during the study period. An efficient filtering method should be able to preserve the variation of its ensemble to properly span the error sub-space. Error covariance matrices are analyzed in terms of estimated errors and correlations.

One important aspect of a filter performance refers to its ability to sample representative ensembles (or particles) at the analysis steps. Figure 8 outlines how the different filtering techniques adjust the ensemble members during the assimilation procedure. The average TWS variations time series over Australia and their ensembles at the analysis steps are calculated for all filters (Figure 8).

FIGURE 8

Several important points can be made from the evolution of ensembles in the assimilation period (Figure 8). Firstly, most of the filters generate ensembles mean (red lines) close to the assimilated observations suggesting that the filters provide good estimates of the observed variables. However, one should also consider the distribution of the ensemble members. Those of EnKF, SQRA, ETKF, EnSRF, and PFSR are consistent over time, which suggests the robustness of these techniques over time. The ensemble members, especially those of the EnSRF and SQRA, are evenly distributed around the mean, implying a good coverage of the error sub-space. The ensembles distribution for DEnKF and EnOI, on the other hand, are different and exhibit an excessively large spread. In most of the cases, the range of the ensemble concentration in DEnKF and EnOI are either misplaced or overestimated. This would result in underestimating the forecast error and possibly inaccurate assimilation results. In the case of PF, the Systematic Resampling technique seems to be more robust; the PFMR ensembles and their variation (black dashed lines) span an unrealistically wide range space, even though the mean appears fairly close to the observations.

More information can be inferred about the filters ensemble distributions by evaluating the ensembles skewness and kurtosis. These indicate the departure of the ensembles distributions from a Gaussian distribution (with a skewness 0 and a kurtosis 3). Kurtosis quantifies the distribution shape (i.e., heavy-tailed or light-tailed, in comparison to a normal distribution) and skewness measures the distribution asymmetry (Joanes and Gill, 1998). The average (forecast and analysis) ensembles skewness and kurtosis of all filters (Figure 9) show skewness and kurtosis are reduced after analysis steps for all filters, suggesting that the filters posterior become closer to Gaussian as assimilation proceeds. This is, however, more pronounced for skewness than for kurtosis, showing the filters higher impact on the ensembles distribution asymmetry. The stochastic EnKF ensemble is closer to a Gaussian distribution, which is related to the application of random noises to the observation (Hoteit et al., 2015). In contrast, the DEnKF and EnOI ensembles are not uniformly distributed, showing a remarkable departure from Gaussian distributions that is expected to introduce bias in the assimilation results.

FIGURE 9

As another evaluation of the filter performance, we further investigate how the model state error covariance changes over time for each filter. The forecast and analysis error covariance matrices at the analysis step indicate how errors change over time, especially after assimilation. We perform two analyses to investigate the influence of the filtering methods on the forecast and analysis error statistics. First, the reductions of error (diagonal elements) in the analysis covariance matrices in comparison to the forecast covariance matrices are calculated at each assimilation step. Next, their minimum, maximum, and average are calculated. The results show how different methods can decrease the errors using GRACE data (Table 5). All the filters reduce errors, where the best performance resulting from SQRA, EnSRF, and, to a less degree, PFSR. Again, DEnKF and EnOI show the weakest effects on error covariance.

TABLE 5

Further insights can be derived from the correlation between the estimated states on the grid points of the study area. For this, 794 grid-points over Australia are considered and the spatial correlation coefficients are computed between each of them and the rest of the points in the assimilation steps. In most of the cases (95%), data assimilation significantly decrease the correlation between grid points in the analysis error covariance matrices. As an example, an arbitrary point approximately in the middle of the study area (for a better visual representation) at the location 136.6854°E and 23.9015°S is chosen and its spatial correlation with the other grid points are plotted to show this effect. The average of spatial correlation map for all assimilation steps and for each filter is presented in Figure 10. Similar results are achieved for the other grid points.

FIGURE 10

One can see from Figure 10 that each filtering method affects the correlation between the specific point and the others differently where some filters like PFs show higher ability to decrease the correlations between errors. This can be related to the native of the algorithm of PF, which produces random particles that are consistent with model nonlinear dynamics. The results of the correlation analysis (cf. Figure 10) are consistent with the other results, with DEnKF and EnOI showing the less ability to reduce errors, also having the least influence on the correlations. These results, along with the outcomes of the ensemble distribution analysis (cf. Figures 8 and 9), demonstrate the effect of successful ensemble generation on estimated errors. The filters (e.g., EnKF, SQRA, and EnSRF) with the higher ability to sample representative ensembles lead to the less estimation errors as well as correlations in contrast to the other filters, especially DEnKF and EnOI.

Only a few filters show a good performance in both analyses. These filters, SQRA, and EnSRF, not only improve the model state estimates compared to GRACE data and the (groundwater level and soil moisture) in-situ measurements but also efficiently decrease the ensemble spread and spatial correlation errors. The resulting estimates of groundwater storage further exhibit less RMSE against independent groundwater level in-situ data.

5. Summary and Conclusions

There is evidence that different filter types are more suited to different applications (Reichle et al., 2002). This study considered the implementation of different data assimilation filtering techniques based on the two most commonly applied algorithms, ensemble Kalman, and Particle Filter, to assess their performances for assimilating GRACE data into the hydrological model of W3RA. GRACE-derived TWS over Australia was assimilated into the W3RA hydrological model using the various filters. Among the ensemble Kalman filters, we tested the stochastic and the deterministic schemes (EnSRF, ETKF, SQRA, DEnKF, and EnOI) along with two different resampling approaches of Particle Filters (PFMR and PFSR). The effects of the filtering methods on the ensembles spread and the estimation error covariance matrices were investigated. The most promising results are obtained using SQRA, EnSRF, and EnKF, both in terms of ensemble generation as well as in dealing with the estimation error covariance matrices. The greatest error reduction with minimum error covariance is achieved by EnSRF (47% average) and SQRA (44% average). These two filters (along with EnKF) also show a good ability to sample representative ensembles with enough spread. The filters state estimates were evaluated against GRACE data, in-situ groundwater measurements, and in-situ soil moisture data. While improvements in the state estimations are observed for all implemented filters, the best results are obtained with, respectively, SQRA (75% correlation to the groundwater level in-situ measurements and 82% correlation to OzNet soil moisture network), EnSRF (42% error reduction), PFSR (37% error reduction) and slightly less successful ETKF (33% error reduction). In contrast, DEnKF was the least successful in dealing with error covariance matrices and suggested a larger error in the state estimates. SQRA and EnSRF, which efficiently dealt with the error covariances, provided the least RMSEs (32.14 mm and 33.74 mm) and maximum correlations to both groundwater level and soil moisture in-situ measurements. These two filters demonstrated a high capability in assimilating GRACE data. GRACE TWS fields are unique in term of resolution, both spatially (almost 3 times rougher than the model) and temporally (monthly). The weak spatial resolution also affects the observation error covariance structure by increasing the correlation between neighboring grid points when working with a fine (e.g., 1°×1°) grid. Therefore assimilating such a dataset could be challenging requiring a filter that is robust to the system error covariances and also powerful in term of resampling representative ensembles after every assimilation step. However, a general conclusion on the preference of ensemble filters might not be possible from this study due to model-specific and application-specific characteristics. Thus, further research might be undertaken to investigate various aspects of filters in different hydrological applications and to explore other filters like new designed PFs that were not considered here.

References

References

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

Altaf, M.U., Butler, T., Mayo, T., Luo, X., Dawson, C., Heemink, A.W., Hoteit, I., (2014). A Comparison of Ensemble Kalman Filters for Storm Surge Assimilation, Monthly Weather Review, 142:8, 2899-2914.

Anderson, J., Anderson, S., (1999). A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. Mon. Weather Rev. 127, 27412758.

AAnderson, J., (2001). An Ensemble Adjustment Kalman Filter for Data Assimilation. Mon. Wea. Rev., 129, 28842903, http://dx.doi.org/10.1175/1520-0493(2001)129¡2884:AEAKFF¿2.0.CO;2.

Anderson, J. L., (2006). Exploring the need for localization in ensemble data assimilation using a hierarchical ensemble filter, Physica D, 230, 99111.

Anderson, M.C., Norman, J.M., Mecikalski, J.R., Otkin, J.A., Kustas, W.P., (2007). A climatological study of evapotranspiration and moisture stress across the continental United States based on thermal remote sensing: 1. Model formulation. J. Geophys. Res. 112 (D10117). http://dx.doi.org/10.1029/2006JD007506.

Arulampalam, M.S., Maskell, S., Gordon, N., Clapp, T., (2002). A tutorial on particle filters for online nonlinear/non-Gaussian Baysian tracking, IEEE Trans. Signal Processes, 50(2), 174188.

Awange, J., Sharifi, M., Baur, O., Keller, W., Featherstone, W., Kuhn, M., (2009). GRACE Hydrological Monitoring of Australia: Current Limitations and Future Prospects. Journal of Spatial Science. 54 (1): pp. 23-35.

Bai, F., (2014). Distributed Particle Filters for Data Assimilation in Simulation of Large Scale Spatial Temporal Systems, Dissertation, Georgia State University.

Bergemann, K., Reich, S., (2010). A mollified ensemble Kalman filter. Q.J.R. Meteorol. Soc., 136: 16361643. http://dx.doi.org/10.1002/qj.672.

Berliner, L.M., Wikle, C.K., (2007). Approximate importance sampling Monte Carlo for data assimilation. Physica D, 230, 3749.

Bertino, L., Evensen G., Wackernagel, H., (2003). Sequential Data Assimilation Techniques in Oceanography, International Statistical Review, Vol. 71, No. 2 (Aug., 2003), pp. 223-241.

Bennett, A.F., (2002); Inverse Modeling of the Ocean and Atmosphere, 234 pp., Cambridge Univ. Press, New York.

Bishop, C. H., Etherton, B., Majumdar, S. J., (2001). Adaptive sampling with the ensemble transform Kalman filter, Part I: theoretical aspects. Mon. Wea. Rev. 129, 420436.

Bocquet, M., Wu, L., Chevallier, F., (2011). Bayesian design of control space for optimal assimilation of observations. Part I: Consistent multiscale formalism. Q.J.R. Meteorol. Soc., 137: 13401356. http://dx.doi.org/10.1002/qj.837.

Brocca, L., Melone, F., Moramarco, T., Wagner, W., Naeimi, V., Bartalis, Z., Hasenauer, S., (2010). Improving runoff prediction through the assimilation of the ASCAT soil moisture product, Hydrol. Earth Syst. Sci., 14, 18811893, http://dx.doi.org/10.5194/hess-14-1881-2010.

Brown, N.J., Tregoning, P., (2010). Quantifying GRACE data contamination effects on hydrological analysis in the MurrayDarling Basin, southeast Australia. Australian Journal of Earth Sciences, 57(3), 329335. http://dx.doi.org/10.1080/08120091003619241.

Burgers, G., van Leeuven, P.J., Evensen, G., (1998). Analysis scheme in the ensemble Kalman filter, Mon. Wea. Rev., 126, 17191724.

Carpenter, J., Clifford, P., Fearnhead, P., (1999). Improved particle filter for nonlinear problems, IEE Proceedings - Radar, Sonar and Navigation, vol. 146, no. 1, pp. 2-7, http://dx.doi.org/10.1049/ip-rsn:19990255.

Chen, J.L., Wilson, C.R., Famiglietti, J.S., Rodell, M., (2007). Attenuation effect on seasonal basin-scale water storage changes from GRACE time-variable gravity. Journal of Geodesy, 81, 4, 237245. http://dx.doi.org/10.1007/s00190-006-0104-2.

Cheng, M.K., Tapley, B.D., (2004). Variations in the Earth's oblateness during the past 28 years. Journal of Geophysical Research, Solid Earth, 109, B09402. http://dx.doi.org/10.1029/2004JB003028.

Chiew, F.H.S., Stewardson, M.J., McMahon, T.A., (1993). Comparison of six rainfall-runoff modelling approaches, J. Hydrol., 147, 136.

Christiansen, L., Krogh, P.E., Bauer-Gottwein, P., Andersen, O.B., Leirio, S., Binning, P.J., Rosbjerg, D., (2007). Local to regional hydrological model calibration for the Okavango River Basin from In-situ and space borne gravity observations. Proceedings of 2nd Space for Hydrology Workshop, Geneva, Switzerland, 12-14.

Clark, M.P., Rupp, D.E., Woods, R.A., Zheng, X., Ibbitt, R.P., Slater, A.G., Uddstrom, M.J., (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(10), http://dx.doi.org/10.1016/j.advwatres.2008.06.005.

Coumou, D., Rahmstorf, S., (2012). A decade of weather extremes Nat. Clim. Change, 2 (7), pp. 16.

Courtier, P., Thpaut, J.N., Hollingsworth, A., (1994). A strategy for operational implementation of 4DVAR, using an incremental approach. Quart. J. Roy. Meteor. Soc., 120,1367-1387.

Counillon, F., Bertino, L., (2009). Ensemble optimal interpolation: Multivariate properties in the Gulf of Mexico. Tellus, 61A, 296308.

Doll, P., Kaspar, F., Lehner, B., (2003). A global hydrological model for deriving water availability indicators: model tuning and validation, J. Hydrol., 270, 105134.

Doucet, A., (1998). On sequential simulation-based methods for Bayesian filtering, Tech. Rep. CUED/F-INFENG/TR 310, Dep. of Eng., Cambridge Univ., Cambridge, UK.

Doucet, A., Freitas, N., Murphy, K., Russell, S., (2000). Rao blackwellised particle filtering for dynamic Bayesian networks, in C. Boutilier and M. Godszmidt (eds), Uncertainty in Artificial Intelligence, Morgan Kaufmann Publishers, pp. 176- 183.

Doucet, A., Freitas, N., Gordon N., (2001). Sequential Monte Carlo methods in practice, Springer-Verlag, New York.

Doucet, A., Cappe, O., Moulines, E., (2005). Comparison of resampling schemes for particle filtering. In 4th International Symposium on Image and Signal Processing and Analysis (ISPA).

Dowd, M., (2006). A sequential Monte Carlo approach to marine ecological prediction. Environmetrics 17, 435455.

Dumedah, G., Coulibaly, P., (2013). Evaluating forecasting performance for data assimilation methods: The ensemble Kalman filter, the particle filter, and the evolutionary-based assimilation, Advances in Water Resources, Volume 60, October 2013, Pages 47-63, ISSN 0309-1708, http://dx.doi.org/dx.doi.org/10.1016/j.advwatres.2013.07.007.

Eicker, A., Schumacher, M., Kusche, J., Dll, P., Mller-Schmied, H., (2014). Calibration/data assimilation approach for integrating GRACE data into the WaterGAP global hydrology model (WGHM) using an ensemble Kalman filter: first results, SurvGeophys, 35(6):12851309. http://dx.doi.org/10.1007/s10712-014-9309-8.

Elbern, H., Schmidt, H., (2001). Ozone episode analysis by fourdimensional variational chemistry data assimilation, J. Geophys. Res., 106, 35693590.

Evensen, G., (1994). Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, J. Geophys. Res., 99, 10, 14310, 162.

Evensen, G., (2003). The ensemble Kalman filter: Theoretical formulation and practical inplementation, Ocean Dynamics, 53, 343367, http://dx.doi.org/10.1007/s10236-003-0036-9.

Evensen, G., (2004). Sampling strategies and square root analysis schemes for the EnKF. Ocean Dyn. 54(6), 539-560.

Evensen, G., (2007). Data Assimilation: The Ensemble Kalman Filter, Springer, 279 pp.

Evensen, G., (2009). Data assimilation. The Ensemble Kalman Filter. Springer, Berlin Heidelberg, 2. edition.

Fairbairn, D, Pring, S. R., Lorenc, A. C., Roulstone, I., (2014). A comparison of 4DVar with ensemble dataassimilation methods. Q. J. R. Meteorol. Soc. 140: 281294. http://dx.doi.org/10.1002/qj.2135.

Forootan, E., Awange, J., Kusche, J., Heck, B., Eicker, A., (2012). Independent patterns of water mass anomalies over Australia from satellite data and models. Journal of Remote Sensing of Environment, Vol.124, Pages 427-443, dx.doi.org/10.1016/j.rse.2012.05.023.

Forootan, E., Didova, O., Schumacher, M, Kusche, J., Elsaka, B., (2014). Comparisons of atmospheric mass variations derived from ECMWF reanalysis and operational fields, over

2003 to 2011. Journal of Geodesy, 88, Pages 503-514, http://dx.doi.org/10.1007/s00190-014-0696-x.

Forootan, E., Khandu, Awange, J., Schumacher, M., Anyah, R., van Dijk, A., Kusche, J., (2016). Quantifying the impacts of ENSO and IOD on rain gauge and remotely sensed precipitation products over Australia. Remote Sensing of Environment,172, Pages 50-66, http://dx.doi.org/10.1016/j.rse.2015.10.027.

Garner, T.W., Wolf, R.A., Spiro, R.W. , Thomsen, M.F., (1999). First attempt at assimilating data to constrain a magnetospheric model, J. Geophys. Res., 104(A11), 2514525152, http://dx.doi.org/ 10.1029/1999JA900274.

Giustarini, L., Matgen, P., Hostache, R., Montanari, M., Plaza, D., Pauwels, V.R.N., De Lannoy, G.J.M., De Keyser, R., Pfister, L., Hoffmann, L., Savenije, H.H.G., (2011). Assimilating SAR-derived water level data into a hydraulic model: a case study, Hydrol. Earth Syst. Sci., 15, 23492365, http://dx.doi.org/10.5194/hess-15-2349-2011.

Gordon, N.J., Salmond, D.J., Smith, A.F.M., (1993). Novel approach to nonlinear/non-Gaussian Bayesian state estimation, IEE Proc. F 140, 107113.

Hamill, T.M., Snyder, C., (2002). Using improved background-error covariances from an ensemble Kalman filter for adaptive observations. Mon Wea Rev 130:15521572. http://dx.doi.org/10.1175/1520-0493(2002)130¡1552:UIBECF¿2.0.CO;2.

Hol, J.D., Schon, T.B., Gustafsson, F., (2006). On Resampling Algorithms for Particle Filters, Nonlinear Statistical Signal Processing Workshop, 2006 IEEE, Cambridge, UK, 2006, pp. 79-82. http://dx.doi.org/10.1109/NSSPW.2006.4378824.

Hoteit, I., Pham, D.T., Blum, J., (2002). A simplified reducedorder kalman filtering and application to altimetric data assimilation in tropical Pacific. J. Mar. Syst., 36, 101127.

Hoteit, I., Triantafyllou, G., Petihakis G., (2005). Efficient data assimilation into a complex, 3-D physical-biogeochemical model using partially-local Kalman filters. Annales Geophysicae, European Geosciences Union, 23 (10), pp.3171-3185.

Hoteit, I., Pham, D.T., Triantafyllou, G., Korres, G., (2008). A new approximate solution of the optimal nonlinear filter for data assimilation in meteorology and oceanography, Monthly Weather Review, 136, 317-334.

Hoteit, I., Luo, X., Pham, D.T., (2012). Particle Kalman Filtering: A Nonlinear Bayesian Framework for Ensemble Kalman Filters, Monthly Weather Review, 140:2, 528-542.

Hoteit, I., Pham, D.T., Gharamti, M. E., Luo, X., (2015). Mitigating Observation Perturbation Sampling Errors in the Stochastic EnKF, Monthly Weather Review, 143:7, 2918-2936.

Houborg, R., Rodell, M., Li, B., Reichle, R.H., Zaitchik, B.F., (2012). Drought indicators based on model-assimilated Gravity Recovery and Climate Experiment (GRACE) terrestrial water storage observations. Water Resour Res 48:W07525. http://dx.doi.org/10.1029/2011WR011291.

Houtekamer, P.L., Mitchell, H.L., (1998). Data assimilation using an ensemble Kalman filter technique, Mon. Wea. Rev., 126, 796 811.

Houtekamer, P.L., Mitchell, H.L., (2001). A Sequential Ensemble Kalman Filter for Atmospheric Data Assimilation, Mon. Wea. Rev., 129:1, 123-137.

Huang, S., Kumar, R., Flrke, M., Yang T., Hundecha, Y., Kraft, P., Gao, C., Gelfan, A., Liersch, S., Lobanova, A., Strauch, M., Ogtrop, F.V., Reinhardt, J., Haberlandt, U., Krysanova, V., (2016). Evaluation of an ensemble of regional hydrological models in 12 large-scale river basins worldwide. Clim Chang. http://dx.doi.org/10.1007/s10584-016-1841-8.

Hunt, B.R., Kalnay, E., Kostelich, E.J., Ott, E., Patil, D.J., (2004). Four-dimensional ensemble Kalman filtering, Tellus 56A, 273277.

Huntington, T.G., (2006). Evidence for intensification of the global water cycle: Review and synthesis, J. Hydrol.,319(14), 8395, http://dx.doi.org/10.1016/j.jhydrol.2005.07.003.

Jardak, M., Navon, I.M., Zupanski M., (2007). Comparison of sequential data assimilation methods for the KuramotoSivashinsky equation, International journal for numerical methods in fluids, Volume 62, Issue 4, 374402, http://dx.doi.org/ 10.1002/fld.2020.

Jazwinski, A.H., (1970). Stochastic Processes and Filtering Theory. Academic Press, 376 pp.

Joanes, A.H., Gill, C.A., (1998). Comparing Measures of Sample Skewness and Kurtosis. The Statistician 47(1): 183189.

Kalman, R. E., (1960). A New Approach to Linear Filtering and Prediction Problems. Transactions of the ASME - Journal of Basic Engineering.

Kalnay, E., (2003). Atmospheric modelling, data assimilation and predictability, Cambridge University Press. pp. xxii 341. ISBNs 0 521 79179 0, 0 521 79629 6. http://dx.doi.org/10.1256/00359000360683511.

Kitagawa, G., (1987). Non-Gaussian state-space modeling of nonstationary time series. J. Amer. Stat. Assoc., 82, 10321063.

Kivman, G.A., (2003). Sequential parameter estimation for stochastic systems, Nonlinear Processes Geophys. 10, 253256.

Klees, R., Revtova, E.A., Gunter, B.C., Ditmar, P., Oudman, E., Winsemius, H.C., (2008). The design of an optimal filter for monthly GRACE gravity models. Geophysical Journal International, 175, 2, 417-432. http://dx.doi.org/10.1111/j.1365-246X.2008.03922.x.

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

Lahoz, W.A., Geer, A.J., Bekki, S., Bormann, N., Ceccherini, S., Elbern, H., Errera, Q., Eskes, H.J., Fonteyn, D., Jackson, D.R., Khattatov, B., (2007). The Assimilation of Envisat data (ASSET) project, Atmos. Chem. Phys., 7, 1773 - 1796.

Lawson, W.G., Hansen, J.A., (2004). Implications of stochastic and deterministic filters as ensemble-based data assimilation methods in varying regimes of error growth. Mon. Wea. Rev., 132: 1966-1981.

Longuevergne, L., Wilson, C.R., Scanlon, B.R., Crtaux, J.F., (2013). GRACE water storage estimates for the Middle East and other regions with significant reservoir and lake storage, Hydrol. Earth Syst. Sci., 17, 48174830, http://dx.doi.org/10.5194/hess-17-4817-2013.

Lyster, P.M., Cohn, S.E., Menard, R., Chang, L.P., Lin, S.J., Olsen, R., (1997). Parallel implementation of a Kalman filter for constituent data assimilation. Mon. Weather Rev., 125, 16741686.

Mayer-Gurr, T., Zehentner, N., Klinger, B., Kvas, A., (2014). ITSG-Grace2014: a new GRACE gravity field release computed in Graz. - in: GRACE Science Team Meeting (GSTM), Potsdam am: 29.09.2014.

McLaughlin, D., (2002). An integrate approach to hydrologic data assimilation: Interpolation, smoothing, and filtering, Adv. Water Resour., 25, 12751286.

Moradkhani, H., Hsu, K.L., Gupta, H., Sorooshian, S., (2005). Uncertainty assessment of hydrologic model states and parameters: Sequential data assimilation using the particle filter, Water Resour. Res., 41, W05012.

Moradkhani, H., DeChant, C.M., and Sorooshian, S., (2012). Evolution of ensemble data assimilation for uncertainty quantification using the particle filter-Markov chain Monte Carlo method, Water Resour. Res., 48, W12520.

Neal, J., Schumann, G., Bates, P., Buytaert, W., Matgen, P., Pappenberger, F., (2009). A data assimilation approach to discharge esti- mation from space, Hydrol. Process., 23, 36413649.

Nerger, L., (2004). Parallel Filter Algorithms for Data Assimilation in Oceanography, PhD Thesis, University of Bremen.

Oke, P. R., Schiller, A., Griffin, D. A., Brassington, G. B., (2005). Ensemble data assimilation for an eddy-resolving ocean model of the Australian Region. Q. Jl R. Met. Soc., 131, 330111.

Oke, P.R., Sakov, P., Corney, S.P., (2007). Impacts of localisation in the EnKF and EnOI: experiments with a small model, Ocean Dyn. 57, 3245.

Oke, P.R., Brassington, G.B., Griffin, D.A., Schiller, A., (2008). The Bluelink Ocean Data Assimilation System (BODAS). Ocean Modelling, 21, 4670, http://dx.doi.org/10.1016/j.ocemod.2007.11.002.

Ott, E., Hunt, B.R., Szunyogh, I., Zimin, A.V., Kostelich, E.J., Corazza, M., Kalnay, E., Patil, D.J., Yorke, J.A., (2004). A local ensemble Kalman Filter for atmospheric data assimilation. Tellus, 56A: 415-428.

Pham, D.T., (2001). Stochastic methods for sequential data assimilation in strongly nonlinear systems, Mon Weather Rev 129: 11941207.

Plaza, D.A., Keyser, R., Lannoy, G.J.M., Giustarini, L., Matgen, P., Pauwels, V.R.N., (2012). The importance of parameter resampling for soil moisture data assimilation into hydrologic models using the particle filter, Hydrol. Earth Syst. Sci., 16(2), 375390.

Reager, J.T., Thomas, A.C., Sproles, E.A., Rodell, M., Beaudoing, H.K., Li, B., Famiglietti, J.S., (2015). Assimilation of GRACE Terrestrial Water Storage Observations into a Land

Surface Model for the Assessment of Regional Flood Potential. Remote Sens. 2015, 7, 14663-14679.

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

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

Robert, C., Blayo, E., Verron, J., (2006). Comparison of reduced-order, sequential and variational data assimilation methods in the tropical Pacific Ocean. Ocean Dynamics 56: 624. http://dx.doi.org/10.1007/s10236-006-0079-9.

Rodell, M., Chen, J., Kato, H., Famiglietti, J.S., Nigro, J., Wilson, C.R., (2007). Estimating groundwater storage changes in the Mississippi River basin (USA) using GRACE, Hydrogeol. J., 15, 159166.

Sakov, P., Oke, P.R., (2008). A deterministic formulation of the ensemble Kalman filter: an alternative to ensemble square root filters, Tellus 60A, 361371.

Schumacher, M., Eicker, A., Kusche, J., Mller, H., Dll, P., (2015). Covariance analysis and sensitivity studies for GRACE assimilation into WGHM, IAG Symp 143:6 pages http://dx.doi.org/10.1007/1345-2015-119.

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

Schunk, R.W., Scherliess, L., Sojka, J.J., Thompson, D.C., (2004). USU global ionospheric data assimilation models, Atmospheric and Environmental Remote Sensing Data Processing and Utilization: an End-to-End System Perspective, (ed. H.-L. A. Huang and H. J. Bloom), Proc. of SPIE, 5548, http://dx.doi.org/10.1117/12.562448, 327-336.

Seoane, L., Ramillien, G., Frappart, F., Leblanc, M., (2013). Regional GRACE-based estimates of water mass variations over Australia: validation and interpretation, Hydrol. Earth Syst. Sci., 17, 4925-4939, http://dx.doi.org/10.5194/hess-17-4925-2013.

Sheffield, J., Goteti, G., Wood, E. F., (2006). Development of a 50-yearhigh-resolution global dataset of meteorological forcings for land surfacemodeling, J. Clim., 19(13), 30883111.

Silverman, B.W., (1986). Density estimation for statistics and data analysis, Chapman and Hall, London. 175 pp.

Snyder, C., Bengtsson, T., Bickel, P., Anderson, J., (2008). Obstacles to high-dimensional particle filtering, Mon. Wea. Rev., 136, 4629 4640.

Smith, P.J., Beven, K.J., Tawn, J.A., (2008). Detection of structural inadequacy in process-based hydrological models: A particle-filtering approach, WATER RESOURCES RESEARCH, VOL. 44, W01410, http://dx.doi.org/10.1029/2006WR005205.

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

Subramanian, A.C., Hoteit, I., Cornuelle, B., Miller, A.J., Song, H., (2012). Linear versus Nonlinear Filtering with Scale-Selective Corrections for Balanced Dynamics in a Simple Atmospheric Model, Atmospheric Sciences, 69:11, 3405-3419.

Sun, Y.A., Morris, A., Mohanty, S., (2009). Comparison of deterministic ensemble Kalman filters for assimilating hydrogeological data, Advances in Water Resources, Volume 32, Issue 2, Pages 280-292, ISSN 0309-1708, http://dx.doi.org/10.1016/j.advwatres.2008.11.006.

Swenson, S., Wahr, J., (2002). Methods for inferring regional surface-mass anomalies from Gravity Recovery and Climate Experiment (GRACE) measurements of time-variable gravity. Journal of Geophysical research, 107, B9, 2193. http://dx.doi.org/10.1029/2001JB000576.

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

Swenson, S., Chambers, D., Wahr, J., (2008). Estimating geocentervariations from a combination of GRACE and ocean model output. Journal of Geophysical research, 113, B08410, http://dx.doi.org/10.1029/2007JB005338.

Talagrand, O., Courtier, P., (1987). Variational assimilation of meteorological observations with the adjoint vorticity equation-Part 1, Theory Q J R Meteorol Soc 113:13111328.

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, Hydrol. Earth Syst. Sci., 19, 2079-2100, http://dx.doi.org/10.5194/hess-19-2079-2015.

Tapley, B.D., Bettadpur, S., Watkins, M., Reigber, C., (2004). The gravity recovery and climate experiment: mission overview and early results. Geophys Res Lett 31:L09607, http://dx.doi.org/10.1029/2004GL019920.

Thomas, A.C., Reager, J.T., Famiglietti, J.S., Rodell, M., (2014). A GRACE-based water storage deficit approach for hydrological drought characterization. Geophys. Res. Lett. 41, 15371545.

Tippett, M.K., Anderson, J.L., Bishop, C.H., Hamill, T.M., Whitaker, J.S., (2003). Ensemble square root filters, Mon. Weath. Rev., 131, 148590.

Tregoning, P., McClusky, S., van Dijk, A.I.J.M., Crosbie, R.S., Pea-Arancibia, J.L., (2012). Assessment of GRACE Satellites for Groundwater Estimation in Australia, National Water Commission, Canberra, 82 pp.

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

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

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

van Dijk, A.I.J.M., Renzullo, L.J., Wada, Y., Tregoning, P., (2014). A global water cycle reanalysis (20032012) merging satellite gravimetry and altimetry observations with a hydrological multi-model ensemble. Hydrol Earth Syst Sci 18:29552973. http://dx.doi.org/10.5194/hess-18-2955-2014.

van Leeuwen, P.J., Evensen, G., (1996). Data assimilation and inverse methods in terms of a probabilistic formulation. Monthly Weather Review 124, 28982913.

Verlaan, M., (1998). Efficient Kalman filtering algorithms for hydrodynamic models, Ph. D. thesis, TU Delft, Delft, http://ta.twi.tudelft.nl/users/verlaan/artikelen/thesis.ps.gz.

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

Wang, X., Bishop, C.H., (2003). A comparison of breeding and ensemble transform Kalman filter ensemble forecast schemes. J. Atmos. Sci., 60, 11401158.

Wahr, J., Molenaar, M., (1998). Time variability of the Earth's gravity field' Hydrological and oceanic effects and their possible detection using GRACE. Journal of Geophysical research, 103, B12, 30, 205-30, 229, http://dx.doi.org/10.1029/98JB02844.

Weerts, A. H., 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, 117, http://dx.doi.org/10.1029/2005WR004093 W09403.

Weerts, A.H., El Serafy, G.Y.H., Hummel, S., Dhondia, J., Gerritsen, H., (2010). Application of generic data assimilation tools (DATools) for flood forecasting purposes, Comput. Geosci. 36, 4, 453-463, http://dx.doi.org/dx.doi.org/10.1016/j.cageo.2009.07.009.

Wahr, J.M., Molenaar, M., Bryan, F., (1998). Time variability of the Earths gravity field: hydrological and oceanic effects and their possible detection using GRACE. J Geophys Res 108(B12):3020530229, http://dx.doi.org/10.1029/98JB02844.

Whitaker, J.S., Hamill, T.M., (2002). Ensemble data assimilation without perturbed observations, Mon. Wea. Rev., 130, 1913 1924.

Wooldridge, S.A., Kalma, J.D., (2001). Regional-scale hydrological modelling using multiple-parameter landscape zones and a quasi-distributed water balance model. Hydrological Earth System Sciences. 5: 59-74.

Xie, L., Liu, B., Peng, S., (2010). Application of scale-selective data assimilation to tropical cyclone track simulation, J. Geophys. Res., 115, D17105, http://dx.doi.org/10.1029/2009JD013471.

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. J Hydrometeorol 9(3):535548, http://dx.doi.org/10.1175/2007JHM951.1.

Zhang, Y., Bocquet, M., Mallet, V., Seigneur, C., and Baklanov, A., (2012). Real-time air quality forecasting, Part I: History, techniques, and current status, Atmos. Environ., 60, 632655.

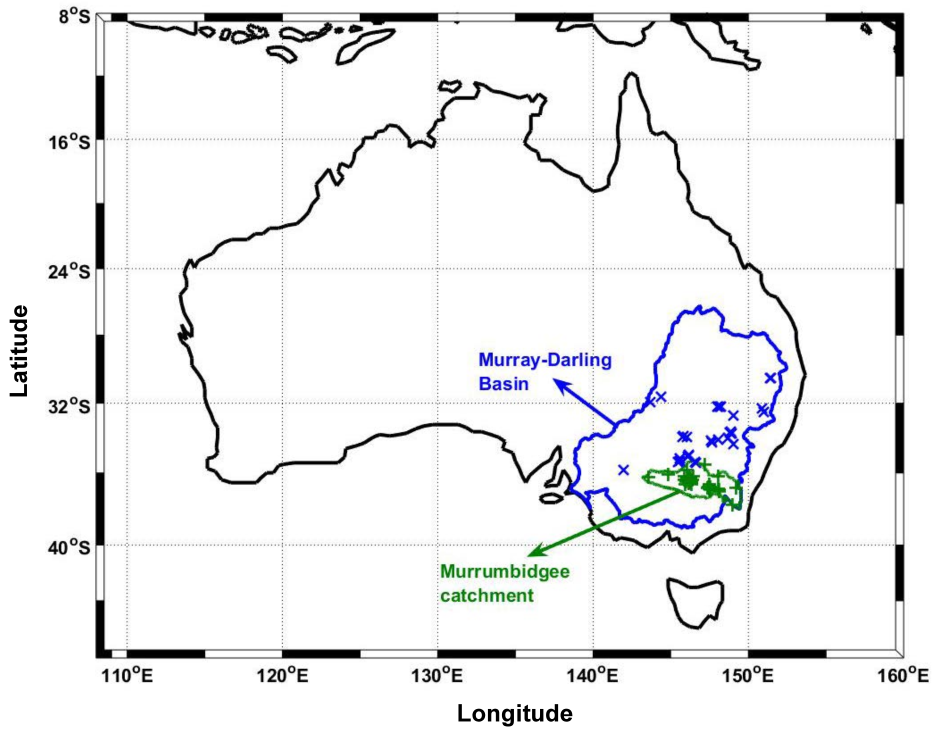

Figure 1: The study area is represented by black solid line. The figure also contains the boundary of the Murray-Darling basin and the locations of the groundwater bore stations (blue), and the outline of the Murrumbidgee catchment with the OzNet soil moisture network (green), which are used for results assessment.

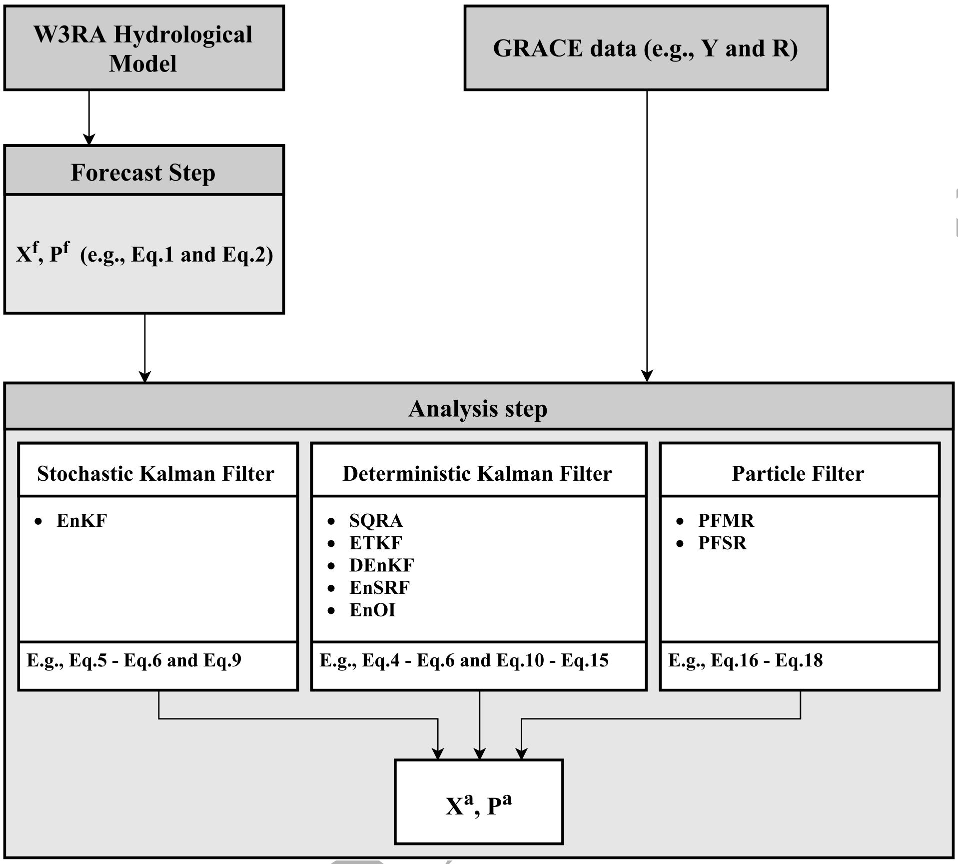

Figure 2: A schematic illustration of the steps and filters applied for data assimilation in this study.

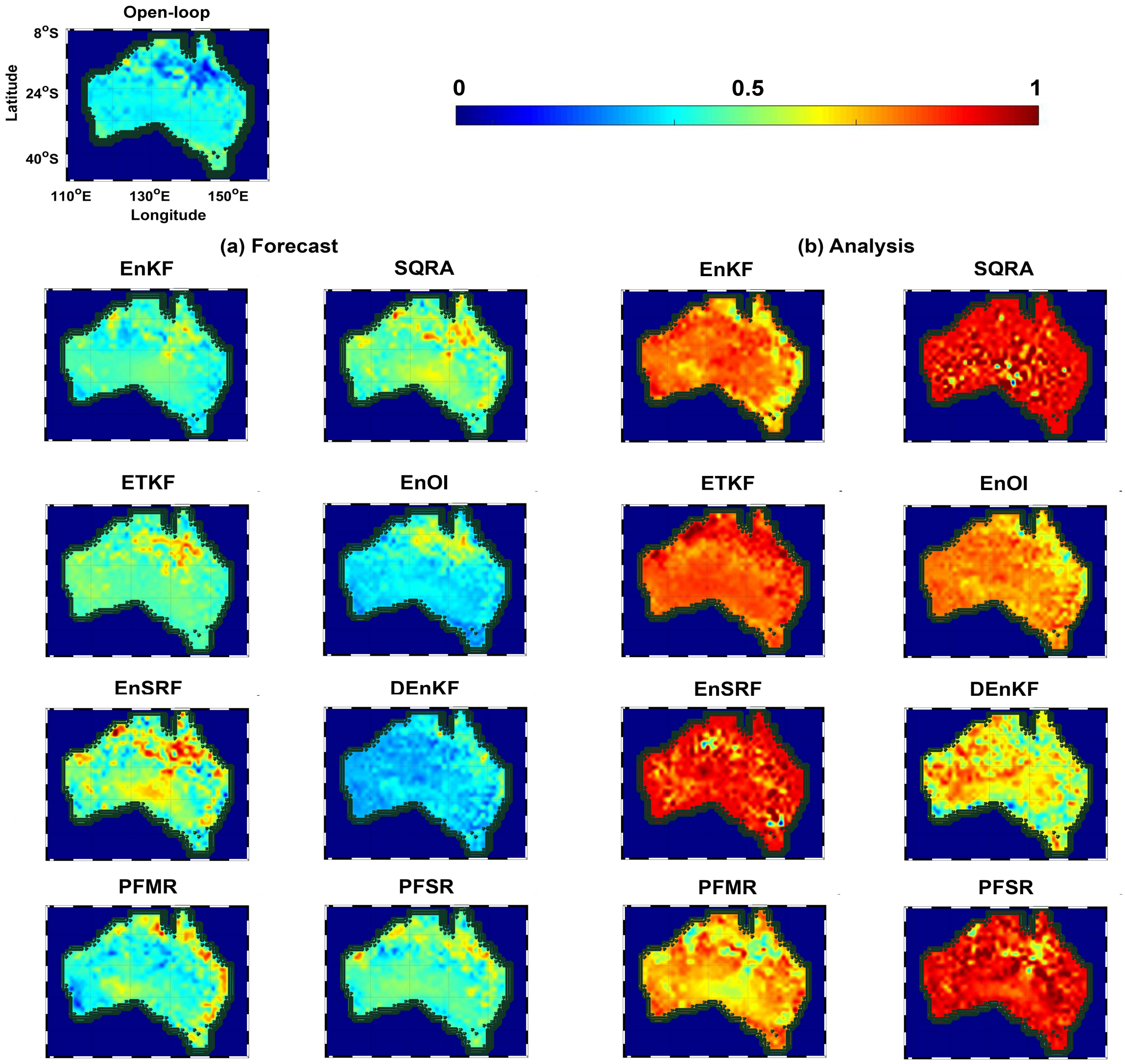

Figure 3: Time average correlations between the assimilated GRACE TWS and different filters estimations at the (a) forecast and (b) analysis steps. Spatial correlation maps are generated at every assimilation step over the study period and their averages are presented.

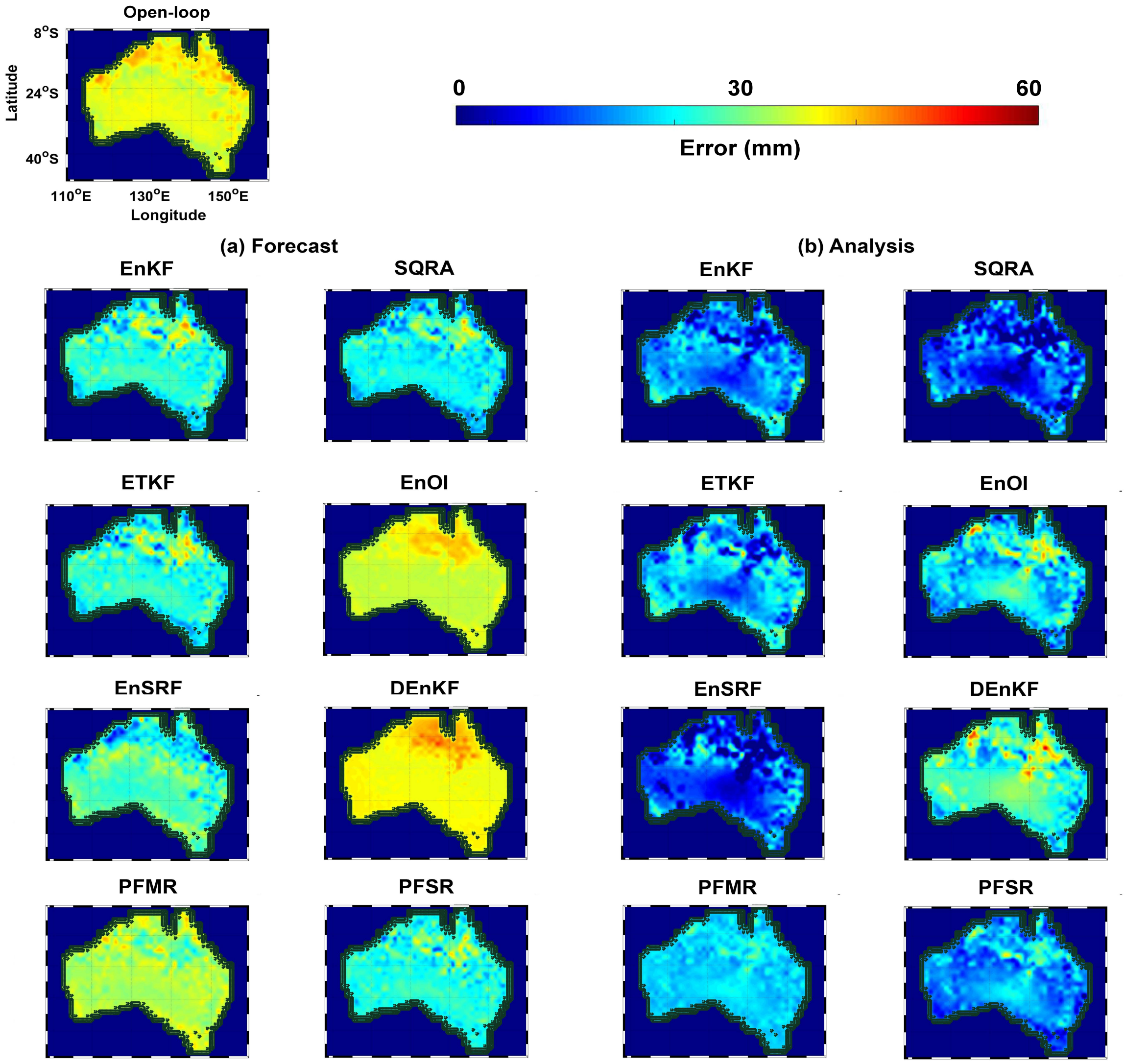

Figure 4: Time average errors between the assimilated GRACE TWS and different filters estimations at the (a) forecast and (b) analysis steps (units are mm). The spatial distribution of the misfits between the filters solution and GRACE data is shown, which plots the time-averaged errors calculated at the forecast steps and the analysis steps.

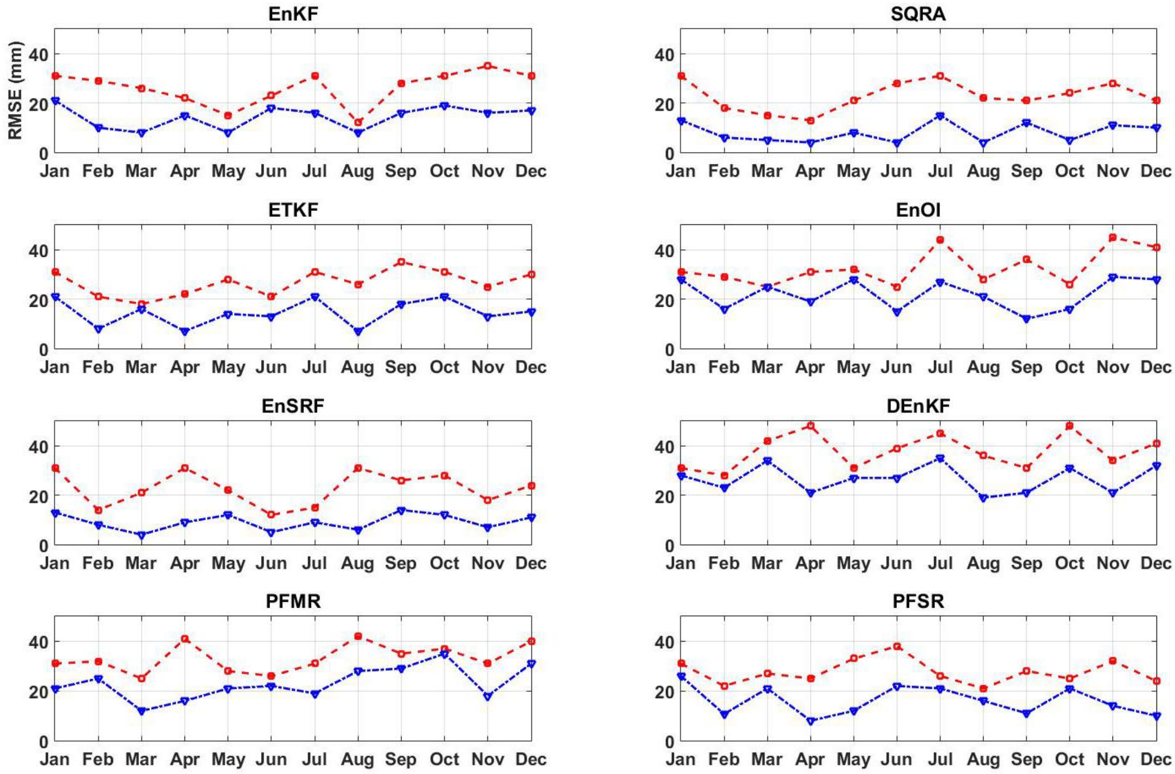

Figure 5: RMSE time series between the assimilated GRACE TWS and the filters' forecasts and analyses which are calculated over all grid points at the forecast (red) and analysis (blue) steps and their averages at each month (during the study period) are shown here.

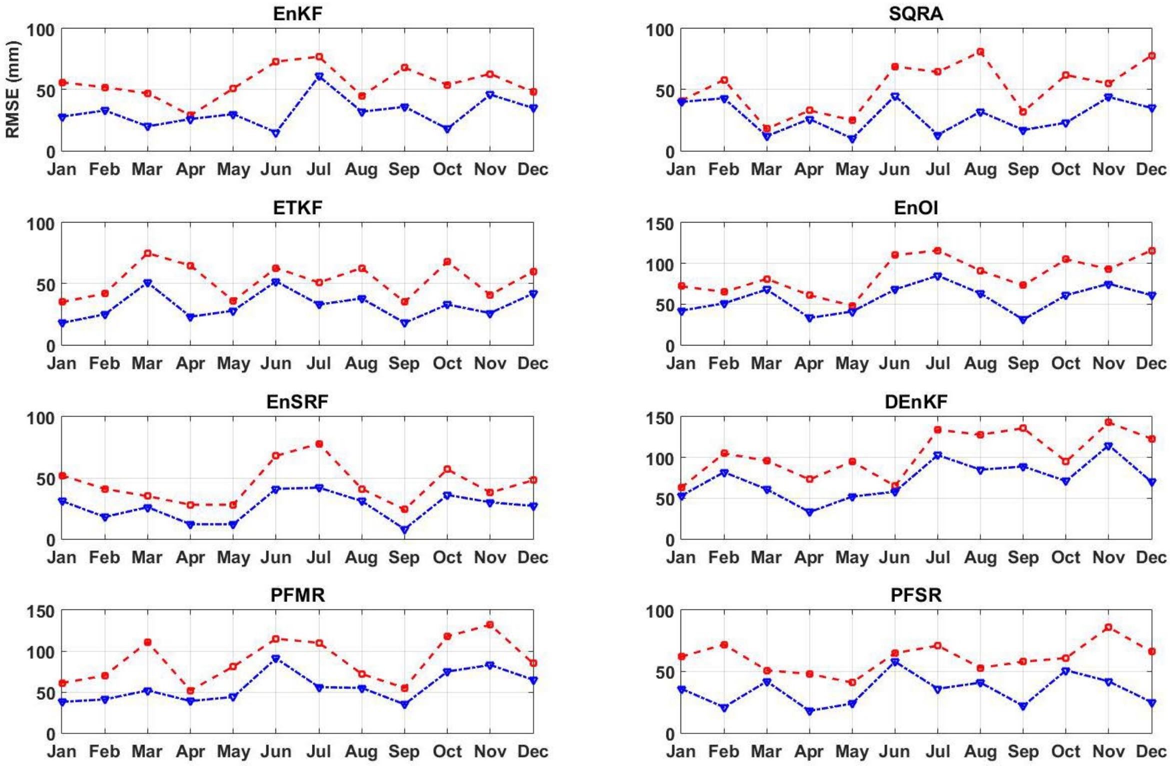

Figure 6: Same as Figure 5, but for the in-situ groundwater measurements and the filters estimates.

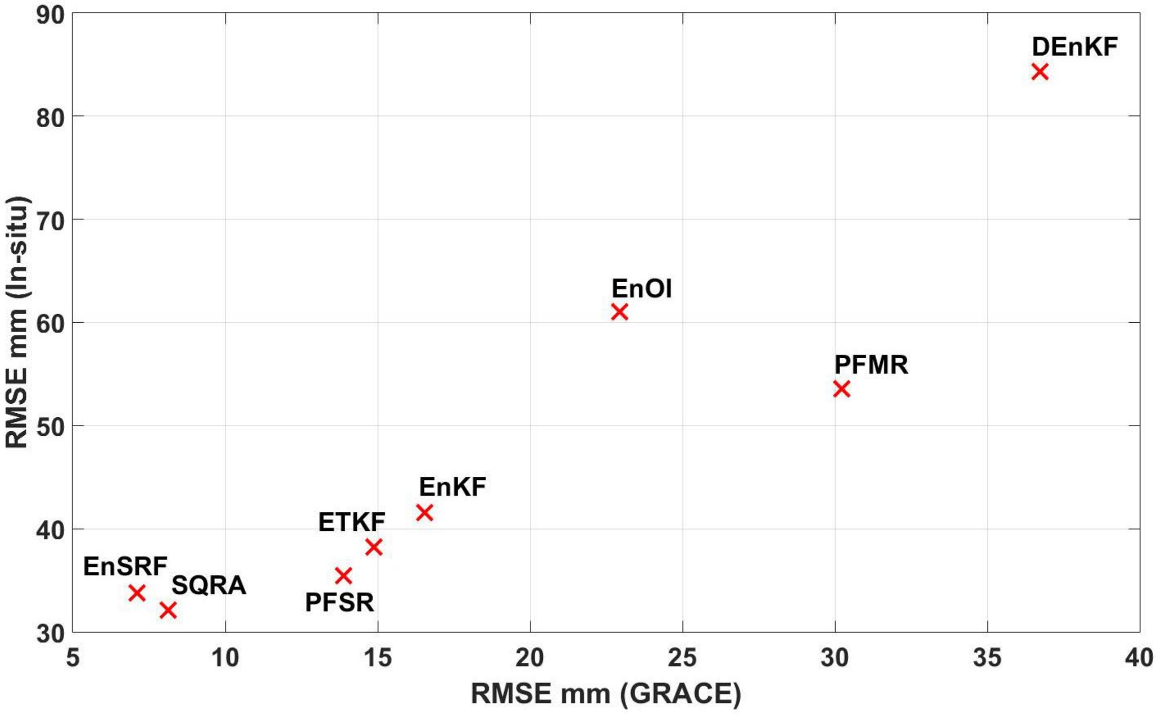

Figure 7: Comparison between the computed average RMSEs of assimilation results from each applied filter using GRACE and the groundwater in-situ datasets. This figure presents the average of the best performances of the filters at the analysis steps from both assessments against GRACE and groundwater in-situs.

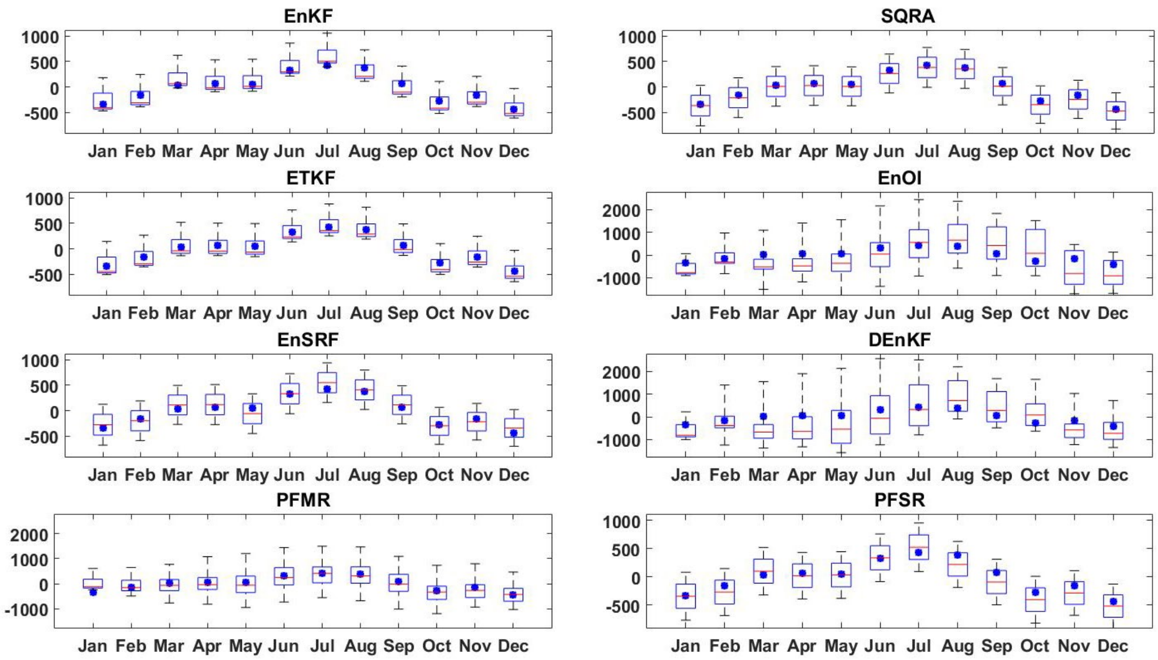

Figure 8: The average TWS variation of ensembles during at the assimilation steps represent by black dashed lines for each filtering method (units are mm). The blue boxes are the ensemble concentrations and horizontal red lines show the median values of the ensembles at each analysis step.

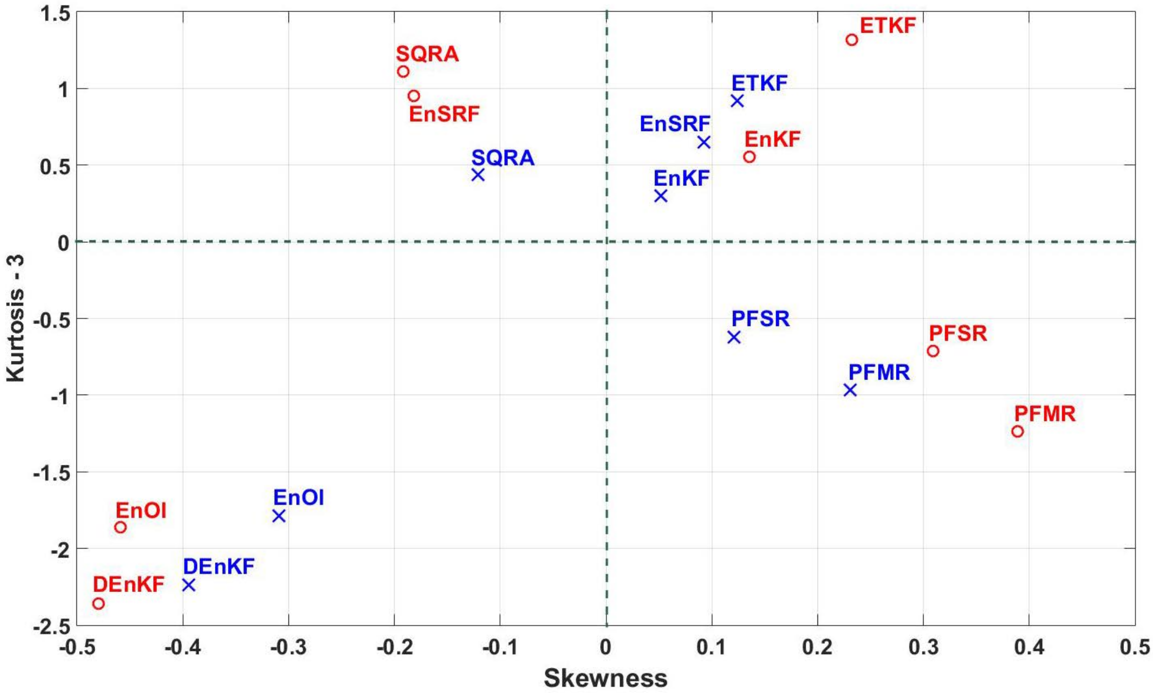

Figure 9: Comparison between the average skewness and kurtosis of each filter for forecast (red circles) and analysis (blue crosses). Note that a normal distribution has a kurtosis of 3 and uses as a reference so the excess kurtosis is usually presented by kurtosis–3.

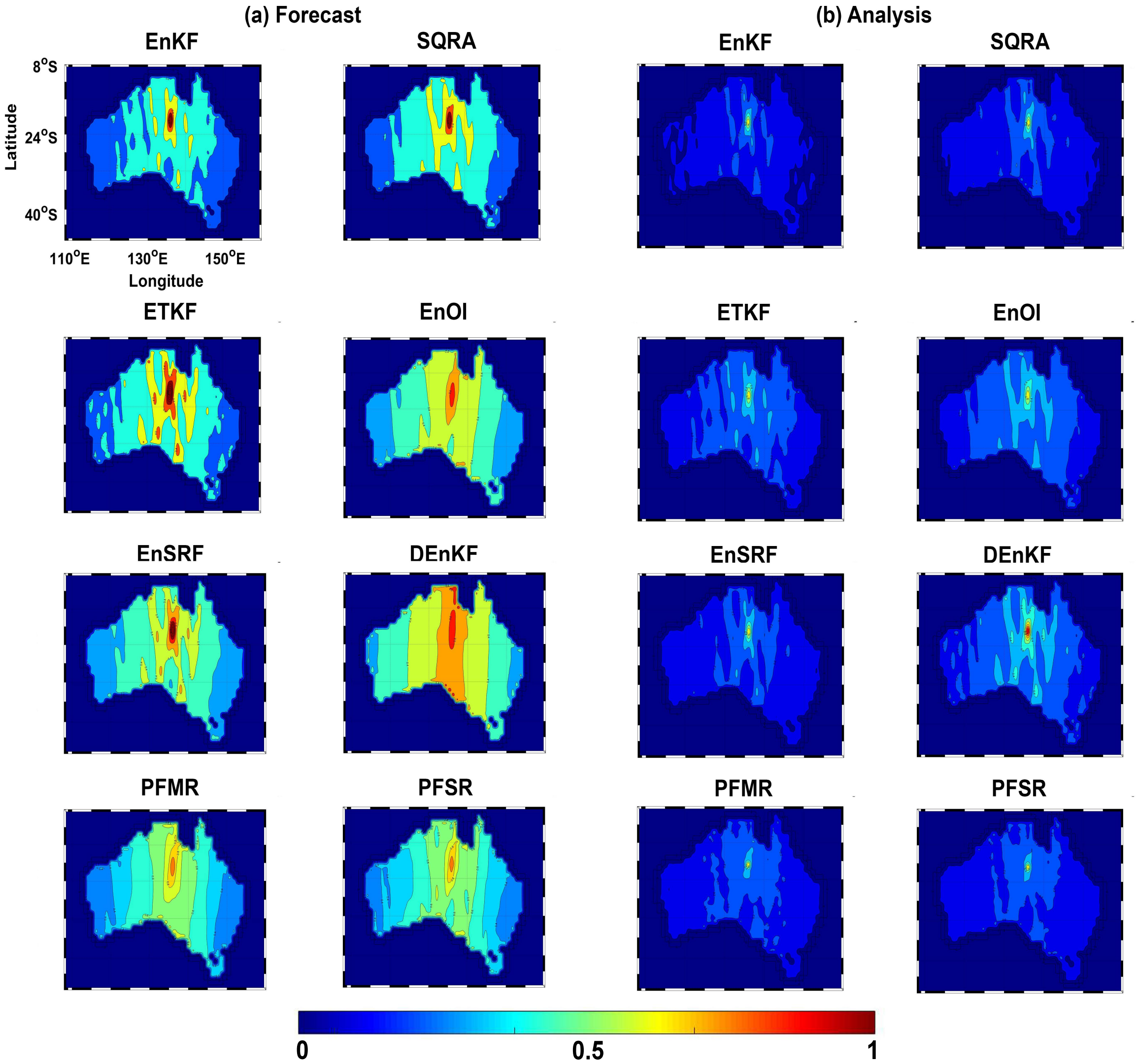

Figure 10: 2-D representation of correlation coefficients of TWS estimated between the arbitrary point (136.6854°E and 23.9015°S) and the rest of the grid points from the covariance matrices. The temporal average of the computed correlation coefficients in forecast and analysis steps are presented.

Table 1: A summary of the applied filters for data assimilation.

Filter	Acronym	Type	Reference
Ensemble Kalman Filter	EnKF	Stochastic ensemble Kalman filter	Evensen (1994)
Square Root Analysis	SQRA	Deterministic ensemble Kalman filter	Evensen (2004)
Ensemble Transform Kalman Filter	ETKF	Deterministic ensemble Kalman filter	Bishop et al. (2001)
Ensemble Square-Root Filter	EnSRF	Deterministic ensemble Kalman filter	Whitaker and Hamill (2002)
Ensemble Optimal Interpolation	EnOI	Deterministic ensemble Kalman filter	Evensen (2003)
Deterministic Ensemble Kalman Filter	DEnKF	Deterministic ensemble Kalman filter	Sakov and Oke (2008)
Particle Filter, Multinomial Resampling	PFMR	Particle filter	Arulampalam et al. (2002)
Particle Filter, Systematic Resampling	PFSR	Particle filter	Arulampalam et al. (2002)

Table 2: A summary of the statistics derived from the implemented methods using the assimilated GRACE data. The improvements in the analysis state RMSE estimates are calculated using the GRACE data in comparison to the model-free run.

	Forecast		Analysis		Improvement (%)
Method	**RMSE** (mm)	R^2	**RMSE** (mm)	R^2	
EnKF	26.5165	0.4354	16.5484	0.9084	39.59
SQRA	18.1156	0.4845	8.1208	0.9335	55.17
ETKF	21.8431	0.4456	14.8704	0.9123	41.92
EnOI	35.2105	0.3951	22.9304	0.7165	34.87
EnSRF	17.2950	0.4912	7.1105	0.9518	58.88
DEnKF	41.6417	0.3610	36.7408	0.6324	15.77
PFMR	37.6009	0.3851	30.2198	0.8137	19.63
PFSR	20.0344	0.4722	13.8711	0.9045	41.74

Table 3: A summary of the statistics derived from implemented methods using the groundwater in-situ measurements. The improvements in the analysis state RMSE estimates are calculated using the in-situ measurements in comparison to the model-free run.

	Forecast		**Analysis**		**Improvement (%)**
Method	**RMSE** (mm)	R^2	**RMSE** (mm)	R^2	
EnKF	62.6521	0.2254	41.5469	0.6456	31.68
SQRA	56.3493	0.2834	32.1387	0.7546	42.96
ETKF	60.7741	0.2574	38.2156	0.6718	33.12
EnOI	89.5411	0.1756	61.0514	0.4675	23.82
EnSRF	58.5271	0.2378	33.7420	0.7225	42.35
DEnKF	112.9712	0.1454	84.3153	0.3385	10.36
PFMR	75.3744	0.1914	53.5445	0.5546	14.96
PFSR	61.0124	0.2246	35.4581	0.6840	37.88

Table 4: A summary of the average correlations between state estimates derived from implemented methods and the soil moisture in-situ measurements. The improvements in the analysis state estimates are calculated using the in-situ measurements in comparison to the model-free run.

Method	**Forecast**	**Analysis**	**Improvement (%)**
EnKF	0.6248	0.7824	25.22
SQRA	0.6524	0.8216	35.93
ETKF	0.6412	0.8003	28.81
EnOI	0.5706	0.6940	21.63
EnSRF	0.6331	0.8431	38.17
DEnKF	0.4867	0.5754	18.22
PFMR	0.5574	0.6835	22.62
PFSR	0.6128	0.7568	32.50

Table 5: Effects of filtering methods on the model state covariance matrix as a percentage improvement.

		Method							
		EnKF	SQRA	ETKF	EnOI	EnSRF	DEnKF	PFMR	PFSR
Error reduction (%)	Minimum	29	35	22	15	34	6	18	28
	Maximum	47	52	44	38	55	20	35	48
	Average	35	**44**	33	21	**47**	8	27	**38**