Water Resources Research

RESEARCH ARTICLE
10.1029/2023WR034544

Special Section:
Hydrogeodesy: Understanding changes in water resources using space geodetic observations

Making the Best Use of GRACE, GRACE-FO and SMAP Data Through a Constrained Bayesian Data-Model Integration

Nooshin Mehrnegar[1], Maike Schumacher[1], Thomas Jagdhuber[2,3], and Ehsan Forootan[1]

[1]Geodesy Group, Department of Planning, Aalborg University, Aalborg, Denmark, [2]Microwaves and Radar Institute, German Aerospace Center, Wessling, Germany, [3]Institute of Geography, University of Augsburg, Augsburg, Germany

Key Points:

- A Constrained Bayesian (ConBay) optimization approach is introduced to integrate GRACE/GRACE-FO and SMAP data into a model
- The ConBay integration of GRACE/GRACE-FO and SMAP controls the updates assigned to the surface soil and groundwater compartments
- The ConBay approach considerably reduced the phase shift differences between the model-derived groundwater storage and USGS observation

Correspondence to:
N. Mehrnegar,
nooshinm@plan.aau.dk

Citation:
Mehrnegar, N., Schumacher, M., Jagdhuber, T., & Forootan, E. (2023). Making the best use of GRACE, GRACE-FO and SMAP data through a Constrained Bayesian data-model integration. *Water Resources Research*, *59*, e2023WR034544. https://doi.org/10.1029/2023WR034544

Received 24 JAN 2023
Accepted 11 AUG 2023

Author Contributions:
Conceptualization: Ehsan Forootan
Data curation: Nooshin Mehrnegar, Thomas Jagdhuber
Formal analysis: Nooshin Mehrnegar, Maike Schumacher, Thomas Jagdhuber, Ehsan Forootan
Funding acquisition: Nooshin Mehrnegar, Ehsan Forootan
Investigation: Nooshin Mehrnegar, Maike Schumacher, Thomas Jagdhuber, Ehsan Forootan

Abstract The Gravity Recovery and Climate Experiment (GRACE, 2003–2017) and its Follow-On mission GRACE-FO (2018-now) provide global estimates of the vertically integrated Terrestrial Water Storage Changes (TWSC). Since 2015, the Soil Moisture Active Passive (SMAP) radiometer observes global L-band brightness temperatures, which are sensitive to near-surface soil moisture. In this study, we introduce our newly developed Constrained Bayesian (ConBay) optimization approach to merge the TWSC of GRACE/GRACE-FO along with SMAP soil moisture data into the ∼10 km resolution W3RA water balance model. ConBay is formulated based on two hierarchical multivariate state-space models to (I) separate land hydrology compartments from GRACE/GRACE-FO TWSC, and (II) constrain the estimation of surface soil water storage based on the SMAP data. The numerical implementation is demonstrated over the High Plain (HP) aquifer in the United States between 2015 and 2021. The implementation of ConBay is compared with an unconstrained Bayesian formulation, and our validations are performed against in-situ USGS groundwater level observations and the European Space Agency (ESA)'s Climate Change Initiative (CCI) soil moisture data. Our results indicate that the single GRACE/GRACE-FO assimilation improves particularly the groundwater compartment. Adding SMAP data to the ConBay approach controls the updates assigned to the surface storage compartments. For example, correlation coefficients between the ESA CCI and the ConBay-derived surface soil water storage (0.8) that are considerably higher than those derived from the unconstrained experiment (−0.3) in the North HP. The percentage of updates introduced to the W3RA groundwater storage is also decreased from 64% to 57%.

1. Introduction

Quantifying Terrestrial Water Storage Changes (TWSC) and its individual compartments (e.g., surface water, canopy, snow, surface/subsurface soil water, and groundwater), as well as their spatial-temporal changes is pivotal for understanding climate variations, sustainability and conservation of water resources (Castellazzi et al., 2016; Forootan, Rietbroek, et al., 2014; Scanlon et al., 2012). Knowing the storage states also plays an important role for many applications related to hydro-meteorology such as flood and drought prediction (see e.g., Forootan et al., 2017, 2019; Houborg et al., 2012; Li et al., 2019; Long et al., 2013).

Various Earth Observation (EO) data sets exist that are used to monitor the total signal or a compartment of large scale TWSC (i.e., known as compartments of hydrological cycle). The Gravity Recovery And Climate Experiment (GRACE, 2002–2017) satellite mission (Tapley, Bettadpur, Ries, et al., 2004; Tapley, Bettadpur, Watkins, & Reigber, 2004) and its Follow-On mission (GRACE-FO, 2018–onward, Tapley et al., 2019; Landerer et al., 2020) provide time-variable Earth's gravity fields that contain signals related to different processes such as non-steric sea level changes, TWSC, ice sheet melting, and Post-Glacial Rebound (PGR), with a spatial resolution of few hundred kilometers and a temporal resolution of daily to monthly (Flechtner et al., 2016). An overview of the application of GRACE and GRACE-FO data can be found in J. Chen et al. (2022). Although GRACE/GRACE-FO TWSC fields represent an accurate superposition of water storage changes, separating this integrated signal into its contributors is desirable for many geodynamic and hydro-climatic applications. Moreover, the limited satellite coverage and sensing depths often restrict the reliability of the remote sensing observations (Reichle, 2008).

Advances in the EO technology have shown that surface soil moisture (top soil layer) can be measured by various remote sensing techniques (Bruckler et al., 1988; Du et al., 2000; Petropoulos et al., 2015). For example, surface soil moisture can be estimated based on optical/thermal remote sensing (Gillies & Carlson, 1995; Sandholt et al., 2002) at regular time intervals, and at spatial scales ranging from a few meters to kilometers.

Methodology: Nooshin Mehrnegar, Maike Schumacher, Ehsan Forootan
Project Administration: Ehsan Forootan
Software: Nooshin Mehrnegar
Supervision: Ehsan Forootan
Validation: Nooshin Mehrnegar
Visualization: Nooshin Mehrnegar
Writing – original draft: Nooshin Mehrnegar
Writing – review & editing: Maike Schumacher, Thomas Jagdhuber, Ehsan Forootan

Several studies have also acknowledged that microwave techniques have a high potential for retrieving soil moisture on a regular basis, either from active or passive sensors (see e.g., Entekhabi et al., 2010; Kerr et al., 2010; Zribi & Dechambre, 2003; Zribi et al., 2005). The Sun-Synchronous-Orbit (SSO) satellite microwave missions such as the L-band radiometer instruments, for example, the Soil Moisture and Ocean Salinity (SMOS, Kerr et al., 2012) and the Soil Moisture Active Passive (SMAP, Entekhabi et al., 2010) provide global surface soil moisture products (Chan et al., 2016) at a spatial resolution of $\sim 25 - 36$ km, which is currently down-scaled to 1 km spatial resolution using vegetation and surface temperature (Fang et al., 2013, 2018, 2020) and Sentinel-1A/Sentinel-1B SAR data (Das et al., 2019; Jagdhuber et al., 2019).

In recent years, various studies indicate that by integrating GRACE/GRACE-FO TWSC into hydrological models, one can spatially downscale and vertically dis-aggregate GRACE/GRACE-FO TWSC into its individual surface and sub-surface water storage estimates. The integration of remote sensing data and models lends more realism to such water storage estimates (see, e.g., Andreadis & Lettenmaier, 2006; Eicker et al., 2014; Girotto et al., 2016, 2017; Mehrnegar, Jones, Singer, Schumacher, Bates, et al., 2020; Miro & Famiglietti, 2018; Schumacher et al., 2018; Tangdamrongsub et al., 2018; Zaitchik et al., 2008a). It has also been shown by many studies that assimilation of the remote sensing soil moisture observations improves the model-based surface soil moisture predictions particularly over poorly instrumented areas of the world that lack good quality precipitation data (Blankenship et al., 2016; Lievens et al., 2015; Reichle & Koster, 2005; Ridler et al., 2014; Tangdamrongsub et al., 2020; Xu et al., 2015). However, it has been found that soil moisture assimilation may introduce a negative impact on the groundwater storage estimate (Girotto et al., 2019; Tangdamrongsub et al., 2020; Tian et al., 2017).

Tian et al. (2017), for example, used the Ensemble Kalman Smoother (EnKS, Evensen & Van Leeuwen, 2000) approach to jointly assimilate GRACE TWSC and SMOS soil moisture data in a water balance model over the Australian continent, and compared its performance against the assimilation of either these data sets, separately. A global extension of this study is performed by Tian et al. (2019), who showed after the joint assimilation, one can better estimate the impact of changes in root-zone soil moisture on vegetation vigor. Girotto et al. (2019) investigated whether the multi-sensor assimilation of GRACE and SMOS observations into the Catchment land surface model can improve the estimation of surface soil moisture, shallow soil water storage (5–100 cm), and (unconfined) groundwater levels. Tangdamrongsub et al. (2020) evaluated the benefit of jointly assimilating SMOS and SMAP soil moisture data and GRACE TWSC into the Community Atmosphere and Biosphere Land Exchange (CABLE) land surface model in South-East Australia. These studies, therefore, suggested that the SMOS/SMAP-only assimilation can improve the surface soil moisture estimates but it reduces the accuracy of TWSC and groundwater estimates, while that of GRACE-only assimilation can improve the groundwater estimates but it does not always produce accurate estimates of surface and shallow soil moisture compartments. Therefore, a multi-sensor assimilation of all these observations is considered in this study.

The objective of this study is to investigate the feasibility and benefits of a fully Bayesian integration scheme, known as the ”Constrained Bayesian (ConBay)” optimization approach, to jointly merge the SMAP soil moisture data and GRACE/GRACE-FO TWSC with the output of a large-scale hydrological model. ConBay was proposed by Forootan and Mehrnegar (2022) based on the hierarchical multivariate state-space models (Koller & Friedman, 2009; Rabiner, 1989) (I) between GRACE/GRACE-FO observations and model outputs, and (II) between the second set of remote sensing observation (e.g., GNSS measurements or remote sensing of (top) soil moisture) and its associated water storage changes derived from (I) (e.g., post-glacial rebound or surface soil water storage, respectively). Therefore, ConBay is flexible to merge various EO data with model outputs, which can be considered as a benefit of this technique. ConBay is applied in this study, while the second state-space model, known as the “constraint equation”, is formulated to use SMAP soil moisture data to constrain the update values of surface soil water storage derived from (I). Another key point of using ConBay, as a multi-sensor merger, is that one can apply it in the offline mode (i.e., a sequentially running of the model after each updating step is not required in ConBay unlike most of the EnKF-based implementations). This means that the ConBay can be easily up- and down-scaled, for example, to be applied both globally and regionally and for different spatial and temporal resolutions.

The ConBay approach accepts model-derived storage estimates as its initial input, therefore, it is flexible to work with any kinds of hydrological models. The ConBay of this study is applied as a “data fusion” or “offline merging” technique, which is not necessarily equivalent to the on-line data assimilation implementation as in for example, Zaitchik et al. (2008a); Eicker et al. (2014); Girotto et al. (2016); Tangdamrongsub et al. (2018);

Schumacher et al. (2018). For simplicity, in the entire paper, the terms of DA is used instead of merger to describe the ConBay technique.

To deal with the temporal evolution of hydrological processes, which is the main difference between the ConBay and previous DA techniques, the update of model states within the commonly applied EnKF-based techniques (Girotto et al., 2016, 2017; Khaki et al., 2017; Schumacher et al., 2016, 2018; Van Dijk et al., 2014; Zaitchik et al., 2008b) is based on the error covariance matrices of each assimilation step. Mehrnegar, Jones, Singer, Schumacher, Jagdhuber, et al. (2020); Forootan and Mehrnegar (2022) showed that an effective Markov chain Monte Carlo (MCMC, Geyer, 1991) algorithm can improve the temporal fit (between model outputs and observations) because: (a) MCMC samples from the Probability Density Functions (PDFs) of unknown time variable weights, which are used to update model outputs against observations, and (b) MCMC allows to estimate the unknown temporal dependency between water storage compartments. Through this formulation, the temporal dependencies can vary in time. As a result, the history of the hydrological processes can be updated sequentially by introducing new measurements. This feature is missing in the previous EnKF-based DA implementations.

Compared to other types of Bayesian approaches, such as the Particle Filter (PF) and Particle Smoother (PS, Särkkä, 2013), ConBay provides the ability to deal with high-dimensional fusion tasks such as global hydrological application (e.g., Bain & Crisan, 2008; Snyder et al., 2008). PF and PS use a set of particles (also called samples) to represent the posterior distribution of stochastic processes given noisy and/or partial observations. Since the computational cost of them grows with the number of particles, choosing a specific number of particles in the design of the filter is a key parameter for these methods. This limitation, however, is addressed in ConBay by formulating the MCMC to sample the unknown parameters conditional on the observations and their uncertainty.

ConBay is implemented here within the High Plain (HP) aquifer in the central Unites States (US). HP is a large first ranked aquifer, which is known for its total groundwater withdrawals (see, e.g., Cano et al., 2018; Maupin & Barber, 2005; Scanlon et al., 2012). For many decades, water from HP has been served as main source for agriculture and public use (Maupin & Barber, 2005; Thelin & Heimes, 1987). For this study, the World-Wide Water balance model (W3RA, Van Dijk, 2010) has been selected to provide a *priori* information of water storage changes in ConBay, and our implementation covers the period of January 2015 to December 2021, where both the SMAP observations and GRACE/GRACE-FO TWSC fields are available. The gap between the GRACE and GRACE-FO missions (between 2017 and 2018) is filled using an in-house iterative reconstruction procedure proposed by Forootan et al. (2020).

To assess the impact of the newly introduced constraint, the Dynamic Model Data Averaging (DMDA, Mehrnegar, Jones, Singer, Schumacher, Bates, et al., 2020), which combines the benefits of Kalman Filtering (KF) and Bayesian Model Averaging (BMA), is implemented to merge GRACE/GRACE-FO TWSC and SMAP soil moisture data (simultaneously) with the W3RA model outputs. The implementation of DMDA is equivalent with the EnKF in an offline mode, but unlike EnKF, it does not require to rerun the model after each DA step, thus, it is closer to the ConBay implementation. We also applied an "unconstrained" Bayesian merger known as the 'Markov Chain Monte Carlo-Data assimilation (MCMC-DA, Mehrnegar, Jones, Singer, Schumacher, Jagdhuber et al., 2020), which is a fully Bayesian DA approach to (only) merge GRACE/GRACE-FO TWSC with model outputs. This comparison will evaluate the contribution of SMAP data on the final water storage estimations. In this study, the outputs of the open loop model runs (of TWSC and individual water storage estimates) are shown to represent dynamics of the original model without the addition of DA observations. This comparison is an important indicator to show how well the original model performs and how much the Bayesian DA can change its estimation.

To validate the groundwater storage estimation results, extensive records of in-situ groundwater level data for 813 stations are used, where the required information about the Storage Coefficients (SCs) to convert these groundwater level measurements to groundwater storage estimates is available from, for example, Gutentag et al. (2014); Peterson et al. (2016); Butler et al. (2020); and Scanlon et al. (2021). The European Space Agency (ESA)'s Climate Change Initiative (CCI) Active soil moisture products are used to validate the surface soil storage estimates. The ESA CCI soil moisture algorithm generates consistent, quality-controlled, and long-term soil moisture climate data, which generally agrees well with the spatial and temporal patterns estimated by land surface models and observed in-situ data (Dorigo et al., 2017).

2. Data and Models

2.1. TWSC From GRACE and GRACE-FO

The release six (RL06) GRACE/GRACE-FO Level 2 (L2) products, with the spherical harmonics of maximum degree and order 60, provided by the Center for Space Research (CSR, http://www2.csr.utexas.edu/grace/) are used to compute monthly TWSC fields covering January 2003 to June 2017, and June 2018 to December 2021. In order to generate monthly TWSC from GRACE/GRACE-FO products, recommended corrections are applied. For example, the degree 1 coefficients, which are not observed by GRACE/GRACE-FO, are replaced by those from Swenson et al. (2008) to account for the movement of the mass center of the Earth. Degree 2 and order 0 (C_{20}) coefficients are replaced by more reliable estimates of the Satellite Laser Ranging (SLR) solutions following J. L. Chen et al. (2007). Surface deformations, also known as the Glacial Isostatic Adjustment (GIA), are reduced using the output of the ICE-6G-D(VM5a) GIA model (Argus et al., 2014; Peltier et al., 2015; Richard Peltier et al., 2018), which is one of the most recently published models of the GIA process in the ICE-NG(VMX) sequence from the University of Toronto.

Correlated errors of the potential coefficients, which are caused by an-isotropic spatial sampling of the mission, instrument noise, and temporal aliasing from incomplete reduction of short-term mass variations (Forootan, Didova, et al., 2014), are reduced by applying the DDK3 filter (Kusche et al., 2009). The formulation in Wahr et al. (1998) is used to convert the L2 potential coefficients to $0.1° \times 0.1°$ gridded TWSC (to be consistence with W3RA hydrological model outputs explained in Section 2.3) within the High Plain aquifer ($96°W - 106°W$, $31°N - 44°N$), while covering the period of 2015–2021. The resolution, selected for gridding the GRACE and GRACE-FO data, is much higher than the signal content of these missions, which is around few hundred kilometers. However, since the model grids are of finer spatial resolution, we trust the spatial pattern dictated from the model and the amplitude is adjusted to the gridded observations.

Uncertainties of TWSC are computed by implementing a collocation error estimation, that is the Three Cornered Hat (TCH) method (Awange et al., 2016; Ferreira et al., 2016), using TWSC estimates from the CSR, Jet Propulsion Laboratory (JPL), and GeoForschungsZentrum (GFZ) L2 data. This means that, thought the formulation of ConBay allows one to include the information on the covariance of observations, the implementation of this study only considered the variance of TWSC errors. This choice is done to speed up the sampling of the MCMC optimization. Uncertainties of the GRACE/GRACE-FO TWSC within the HP aquifer is estimated to be 16 mm, on average, using the TCH method.

2.2. Filling the Gaps of TWSC Fields

Forootan et al. (2020) introduced an iterative decomposition approach to reconstruct (i.e., merging the gap of) GRACE and GRACE-FO data. This reconstruction approach uses the TWSC derived from the temporal gravity field products of ESA's Swarm mission (Bezděk et al., 2016) as initial values for the missing fields. Then, the Independent Component Analysis (ICA, Forootan et al., 2012; Forootan & Kusche, 2013) is applied to update these initial values using the statistics existing in the time series of GRACE, GRACE-FO, and Swarm TWSC fields. This updating procedure is first affected by the Swarm fields that are much smoother than GRACE/GRACE-FO data due to differences in the spatial resolution. However, the iteration adjusts the empirical independent components to make a consistent evolution derived from the original GRACE and GRACE-FO time series and those of the updated gap values.

2.3. SMAP Remote Sensing Soil Moisture Data

The monthly average of the enhanced Level-3 (L3) soil moisture products (O'Neill et al., 2021), version 5, retrieved by the SMAP radiometer, are used in this study to constrain the estimation of near-surface soil moisture (top soil layer). SMAP measurements provide direct sensing of soil moisture in the top 5 cm of the soil column from April 2015 to the present. SMAP L3 product is a daily composite of SMAP Level-2 (L2) soil moisture which is derived from SMAP Level-1C (L1C) interpolated brightness temperatures on the 9 km Equal-Area Scalable Earth Grid ($\sim 0.1° \times 0.1°$ grid data). The volumetric units ($m^3\ m^{-3}$) of the SMAP soil moisture is converted to the vertical changes in soil water storage (in mm) using the first layer (0–5 cm) of the STATSGO porosity values (http://www.soilinfo.psu.edu/).

2.4. W3RA Water Balance Model

Outputs of the Worldwide Water Resources Assessment (W3RA, Van Dijk, 2010) are monthly averaged states of (daily model outputs of) snow, surface water storage, surface soil water (top layer), shallow-rooted soil water, deep-rooted soil water storage, and groundwater storage. These are used in this study as a *priori* information of TWSC components. For this study, the original code (http://wald.anu.edu.au/challenges/water/w3-and-ozwald-hydrology-models/) is modified for the HP aquifer. For this, daily averages of $0.1° \times 0.1°$ ERA5-Land hourly fields (Muñoz Sabater et al., 2019) of precipitation, surface solar radiation downwards, albedo, and 10-m wind, as well as interpolated $0.1° \times 0.1°$ fields of minimum and maximum temperature from ERA5 hourly data on single level (Hersbach & Dee, 2016), with the original resolution of $0.25° \times 0.25°$ are used as forcing data.

Model uncertainty is estimated following Renzullo et al. (2014) by using the perturbed meteorological forcing approach, where an additive error is assumed for the short-wave radiation perturbation of 50 Wm^2, a Gaussian multiplicative error of 30% for rainfall perturbation, and a Gaussian additive error of 2°C as the magnitude of the additive error air temperature perturbations.

Our motivation to select W3RA is its simplicity, which makes its computational load manageable for scientific applications, see examples of the W3RA's applications in, e.g., Khaki et al. (2017); Forootan et al. (2019), and its acceptable performance when compared with other commonly used global hydrological or land surface models (Schellekens et al., 2017).

2.5. In-Situ USGS Groundwater Level Data

The groundwater level data in the HP aquifer are collected from the US Geological Survey (USGS) groundwater network (https://water.usgs.gov/ogw/networks.html), which contains a record of groundwater levels between 1970 until now across the Conterminous United States (CONUS). USGS groundwater observations are an independent validation data set to evaluate groundwater storage changes derived from W3RA with and without assimilation of remote sensing data. The point-wise groundwater time series within the HP aquifer are downloaded for the period 2015–2021 and are filtered to exclude measurements with large data gaps (temporal gaps >4 months), and those time series which only contain linear and/or non-linear trends, without any other oscillations are excluded from data sets. The 813 selected groundwater level wells, which are mostly located in the Northern HP, are then temporally averaged and interpolated to produce monthly time series to be comparable with monthly calculated W3RA groundwater estimates before and after assimilation. Moreover, an effective storage coefficient (Sc) is applied to convert groundwater level to groundwater storage, where groundwater storage = Sc × groundwater level. The Sc values within the HP aquifer are derived from the reported values in Gutentag et al. (2014).

2.6. ESA CCI Satellite-Derived Soil Data

The European Space Agency's Climate Change Initiative (ESA CCI) soil moisture product (Gruber et al., 2019) is used in this study to validate the top layer (<5 cm) soil water storage of W3RA before and after assimilation with GRACE/GRACE-FO TWSC and SMAP soil data. The ESA CCI soil moisture algorithm generates long-term soil moisture climate data by harmonizing and merging soil moisture retrievals from multiple satellites into (a) an active-microwave-based only (ACTIVE), (b) a passive-microwave-based only (PASSIVE) and a (c) combined active–passive (COMBINED) product (Dorigo et al., 2017). According to Dorigo et al. (2017) and a review of existing literature, the ESA CCI product quality has steadily increased with each successive release and the merged products generally outperform the single-sensor products.

The ESA CCI PASSIVE product and the ESA CCI COMBINED product contains also SMAP data, but to a low degree (Gruber et al., 2019). In this study, therefore, we use the ACTIVE ESA CCI product, which is statistically independent from SMAP data. It means that SMAP data is not used in ACTIVE ESA CCI products. To this aim, version 07.1 of daily ESA CCI soil moisture with a spatial resolution of $0.25° \times 0.25°$ and covering the period of 1978–2021 is downloaded from the ESA website (http://www.esa-soilmoisture-cci.org). The monthly ESA CCI soil moisture time series between 2015 and 2021 is computed by temporal averaging of the daily products. These values are then spatially interpolated (using linear interpolation, Meijering, 2002) on the same $0.1° \times 0.1°$ grid that is defined for the W3RA simulations. In accordance with the SMAP data, the volumetric units ($m^3\ m^{-3}$) of

the ESA CCI are converted to vertical changes in soil water storage (in mm) using the first layer (0–5 cm) of the STATSGO porosity values (http://www.soilinfo.psu.edu/).

3. Method

Learning dynamical systems (Thelen & Smith, 1998), also known as system identification modeling or time series modeling, aims to create a model or improve an existing model based on measured signals (Lennart, 1999). In this study, the state-space models, also known as hidden Markov models (Rabiner, 1989) or latent process models (Koller & Friedman, 2009), are applied. They describe the probabilistic dependence between the unobserved (latent) state variables (e.g., hydrological model outputs) and the observed measurement (e.g., GRACE/GRACE-FO TWSC and SMAP soil moisture data). The latent states contain the information about the dynamic system (e.g., hydrological processes within the Earth system) and they allow for a succinct representation of the dynamics in the form of a Markov chain within the state-space model.

The state-space model between GRACE/GRACE-FO observation and a priori information (unobserved variables) can be represented by the observation equation (Eq. (1)) and the state equation (Eq (2), Bernstein, 2005). as

$$Y_t = Z_t\Theta_t + X_t\beta_t + \varepsilon_t, \tag{1}$$

$$[\Theta_{t+1}, \beta_{t+1}] = [\Theta_t, \beta_t] + \delta_t. \tag{2}$$

In Equation (1), Y_t represents the vector of GRACE/GRACE-FO observations for P spatial grid points at time $t = 1, 2, \ldots, T$ (i.e., $Y_t = [y_1, y_2, \ldots, y_P]_t$), while Z_t and X_t are two diagonal matrices to store the a priori information from for example, W3RA hydrological model outputs. In this study, the diagonal elements of $Z_{t(P\times P)}$ contain the W3RA-derived surface water, snow, shallow-soil water, and groundwater storage, and the diagonal elements of $X_{t(P\times P)}$ contains the surface soil water storage (top layer) of W3RA for the spatial grid point $p = 1, 2, \ldots, P$ at time t. Each of the diagonal elements of Z_t is a $1 \times K$ vector of $\left[z_{1,p}, z_{2,p}, \ldots, z_{K,p}\right]_t$, where K is the number of individual water storage components. In Equation (1), Θ_t is a $P \times 1$ vector, where each element itself is a $K \times 1$ vector containing the unknown state parameters, that is, $\left[\theta_{1,p}, \theta_{2,p}, \ldots, \theta_{K,p}\right]_t^T$, to make a relationship between the observation and a *priori* information stored in Z_t. The unknown state parameters related to the surface soil water storage (top layer) for $p = 1, 2, \ldots, P$ spatial grid points at time t are shown by the $P \times 1$ vector β_t.

In Equations (1) and (2), ε_t and δ_t are the residuals corresponding to the observation equation and state equation, respectively. The distribution of ε_t is assumed to be Gaussian with the mean value of zero and the error covariance matrix of V_t, which varies over time and reflects the uncertainty of GRACE/GRACE-FO measurements. The state residual δ_t is assumed stationary Gaussian distributed and independent from ε_t, with the mean value of zero and an error covariance matrix of Q. Thus, the distribution of the additive innovations ε_t and δ_t can be written as

$$\varepsilon_t \sim N(0, V_t), \qquad \delta_t \sim N(0, Q), \tag{3}$$

where V_t is the error covariance matrix of observations.

The central hypothesis in formulating the Bayesian optimization algorithm is that the magnitude of changes in water storage components depends on the history of hydrological processes. However, there is no or little physical knowledge about how this dependency varies over time. Therefore, the error covariance matrix Q in Equation (3) defines the temporal dependency between various compartments of the a *priori* information. In practice, Q is a $K \times K$ matrix, where K is the number of hydrological compartments, and is unique for each spatial grid point. The diagonal elements of Q shows temporal dependency to the previous time steps for each TWSC compartments, and the other element shows temporal correlation between various compartments. Entries of Q are unknown and, therefore, they are simultaneously estimated with the unknown state parameters Θ_t and β_t within the ConBay procedure.

The multivariate state-space model needs to define the vertical and horizontal correlations between the different water compartments but this is extremely difficult. If one compartment, such as the groundwater storage, is dominant, it has also a dominant effect on the estimation of the correlations. Thus, it might be required to use further remote sensing data to better constrain the less dominant compartments, such as the surface soil water storage. These concerns can be addressed in the Bayesian fusion technique with specific inequality/equality constraints on the means and regression coefficients to update different compartment of TWSC. To this aim, in

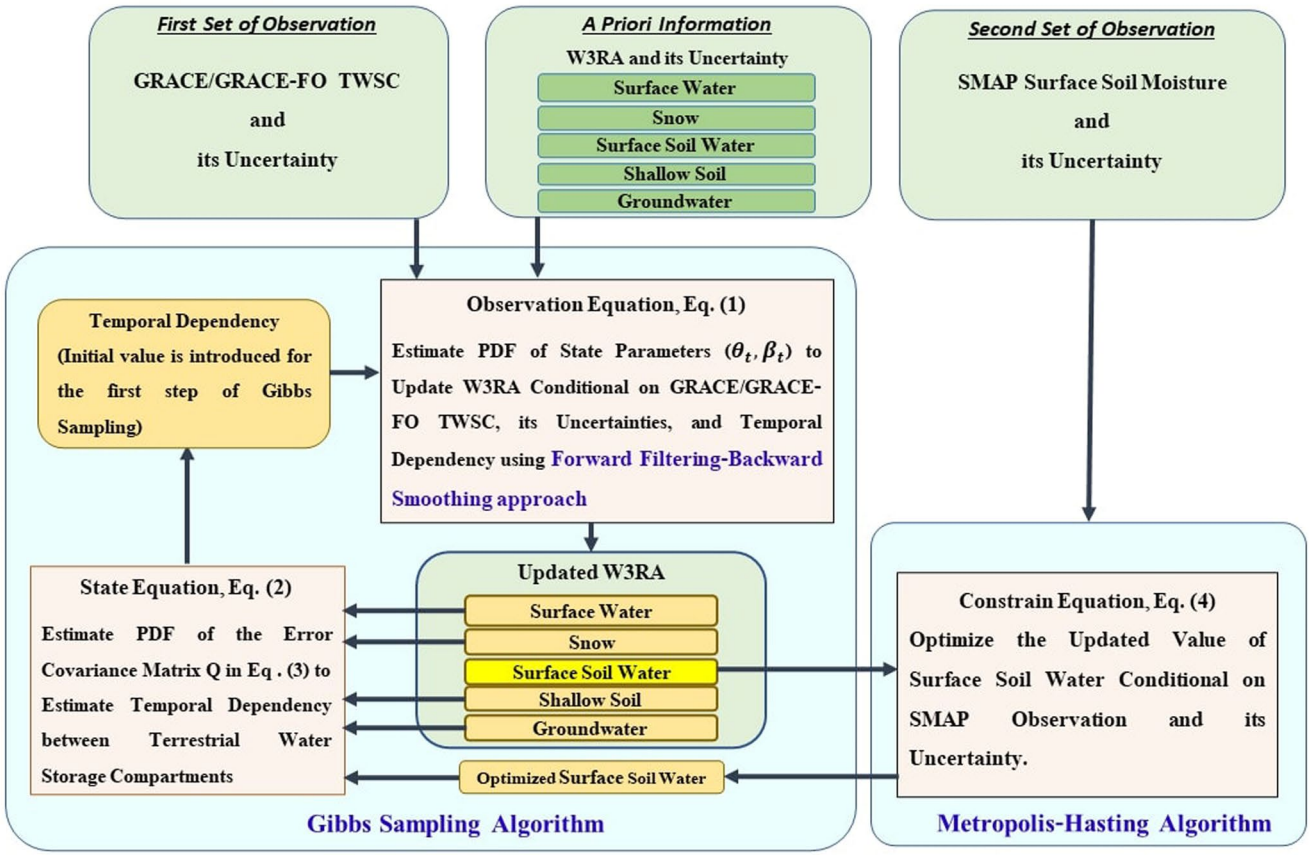

Figure 1. Flowchart of the ConBay method. The framework can accept an arbitrary number of models, and it can be extended to accept various types of observations to merge with the model outputs. PDF in this flowchart stand for the Probability Density Function.

the ConBay approach, a hierarchical constraint equation is formulated to use the second observation data set, for example, SMAP soil moisture observation, for controlling the sampling of β_t derived from Equations (1) and (2) as

$$G_t = X_t\beta_t + \gamma_t, \qquad \gamma_t \sim N(0, U_t), \tag{4}$$

where G_t represent the vector of the SMAP observations, and γ_t is the additive innovation, which is assumed to be Gaussian with the mean value of zero and the error covariance matrix of U_t which contains the uncertainty of the SMAP data. The constraint equation in ConBay is designed to control the updating values of TWSC compartments through an iteration algorithm. In other words, the updated values of W3RA water storage compartments using GRACE/GRACE-FO data is only accepted when the updated values of surface soil moisture ($X_t\beta_t$) is fitted, or close enough, to the SMAP data. In summary, ConBay aims to improve W3RA water balance model in which (a) the summation of all water storage compartments be equal to the GRACE/GRACE-FO TWSC (Equation (1)), and (b) the surface soil water storage be fitted to the SMAP soil moisture data (Equation (4)).

The ConBay algorithm is formulated as a combination of forward-filtering backward-smoothing recursion approach (Kitagawa, 1987), and a Gibbs sampling algorithm (Gelfand & Smith, 1990; Smith & Roberts, 1993) to estimate the unknown state parameters Θ_t and β_t, and the error covariance matrix Q. The generated samples of β_t in each iteration of Gibbs sampling are not accepted automatically as posterior samples; instead they are introduced as candidate samples to a hierarchical Metropolis-Hastings (Chib & Greenberg, 1995) algorithm to be accepted or rejected based on the SMAP measurements. The mathematical formulations of forward-filtering backward-smoothing recursion approach, Gibbs sampling (Gelfand & Smith, 1990; Smith & Roberts, 1993) algorithm, and the hierarchical Metropolis-Hastings algorithm are explained in details in Forootan and Mehrnegar (2022). The work-flow of the ConBay approach is also summarized in Figure 1.

3.1. Departure From the Unconstrained Bayesian Formulations

The MCMC-DA (Mehrnegar, Jones, Singer, Schumacher, Jagdhuber, et al., 2020), which is formulated based on the state-space model (Equations (1) and (2)) is a good candidate of an unconstrained Bayesian approach to separate GRACE/GRACE-FO signals without considering extra information about the surface soil water storage. In fact, MCMC-DA and ConBay work similarly to estimate unknown state parameters and temporal dependency between them, while religiously account for the uncertainties of the observations and a *priori* information. Their main difference is that the hierarchical Metropolis-Hasting algorithm to constrain the surface soil water storage based on SMAP measurements is not implemented in the MCMC-DA. Thus, MCMC-DA is an "unconstrained" Bayesian technique. A comparison between these two techniques will show how beneficial is the joint application of GRACE/GRACE-FO TWSC and SMAP data in signal separation studies.

3.2. Departure From the DMDA Formulation

DMDA (Mehrnegar, Jones, Singer, Schumacher, Bates, et al., 2020) is formulated based on the state-space model (Equations (1), (2)) to merge GRACE/GRACE-FO TWSC with multiple a *priori* information in two steps: (a) a Kalman Filter (Kalman, 1960) approach is applied to solve the state-space model between a set of observation and a *priori* information. (b) The Bayesian Model Averaging (BMA, Hoeting et al., 1999) is considered to provide a time-variable weights to average the water states derived from multiple a *priori* information of the first step, yielding the best fit to the observation. In this study, for the first step we have two sets of a *priori* information for individual water storage compartments: (a) W3RA model outputs, (b) the surface soil water storage derived from SMAP data plus other individual water storage compartments derived from W3RA. The main difference between the DMDA and two other techniques is that, in DMDA, the Kalman Filter approach is formulated to estimate the unknown state parameters, while the unknown temporal dependency between the water states is empirically estimated. All the equations to implement DMDA are explained by details in (Mehrnegar, Jones, Singer, Schumacher, Bates, et al., 2020).

4. Results and Discussion

First, we provide an overview of the differences between GRACE/GRACE-FO TWSC and W3RA TWSC within the HP aquifer between 2015 and 2021. Considerable differences are found between the measured and modeled TWSC in terms of linear trend (Figure 2 (A1) and (B1)) and seasonality, where the annual amplitudes are shown in (Figure 2 (A2) and (B2)). Statistical significance has been tested using the non-parametric Wilcoxon-Mann-Whitney statistical test (WMW, Fay & Proschan, 2010) with 95% levels. TWSC simulated by W3RA shows negative trends with the mean of ~ -10 mm/yr within the HP aquifer between 2015 and 2021, while GRACE/GRACE-FO TWSC shows strong positive trends in the east and northeast (~15 mm/yr) and strong negative trends (~ -20 mm/yr) in the west and southwest of the region. The average of the annual amplitude fitted to the GRACE/GRACE-FO TWSC is estimated to be ~38 mm, while this value is simulated by W3RA to be less than 10 mm in most of the HP region. Further differences between the W3RA outputs and GRACE/GRACE-FO TWSC can be seen in terms of the phase of annual amplitude (computed after removal of linear trends from TWSC), which vary between $\pm 120°$ within the HP aquifer (Figure 2 (B3)).

4.1. TWSC Derived From Bayesian Data Assimilation

The obtained results indicate that after merging W3RA with GRACE/GRACE-FO TWSC, through MCMC-DA, the strong negative and positive trends observed by GRACE/GRACE-FO are introduced to the modeled TWSC (Figure 2 (C1)), and the same spatial pattern of linear trends in the GRACE/GRACE-FO TWSC is obtained. Similar results can be seen in the ConBay output, where we merge GRACE/GRACE-FO and SMAP soil moisture data with W3RA (Figure 2 (D1)). Beside the linear trends, the seasonality of W3RA TWSC has been significantly changed after implementing both Bayesian techniques. For example, the mean of annual amplitude increased from 15 to 40 mm after implementing both MCMC-DA and ConBay approaches (Compare Figure 2 (B2) with Figure 2 (C2) and (D2)). Therefore, as expected, both Bayesian merging techniques performed well to update TWSC of the model, in which the Root Mean Square of Differences (RMSD) between GRACE/GRACE-FO TWSC and W3RA TWSC are decreased from 38 mm to less than 5 mm after implementing both techniques. Another significant changes can be seen in terms of phase of the TWSC, where the phase differences of $\pm 120°$

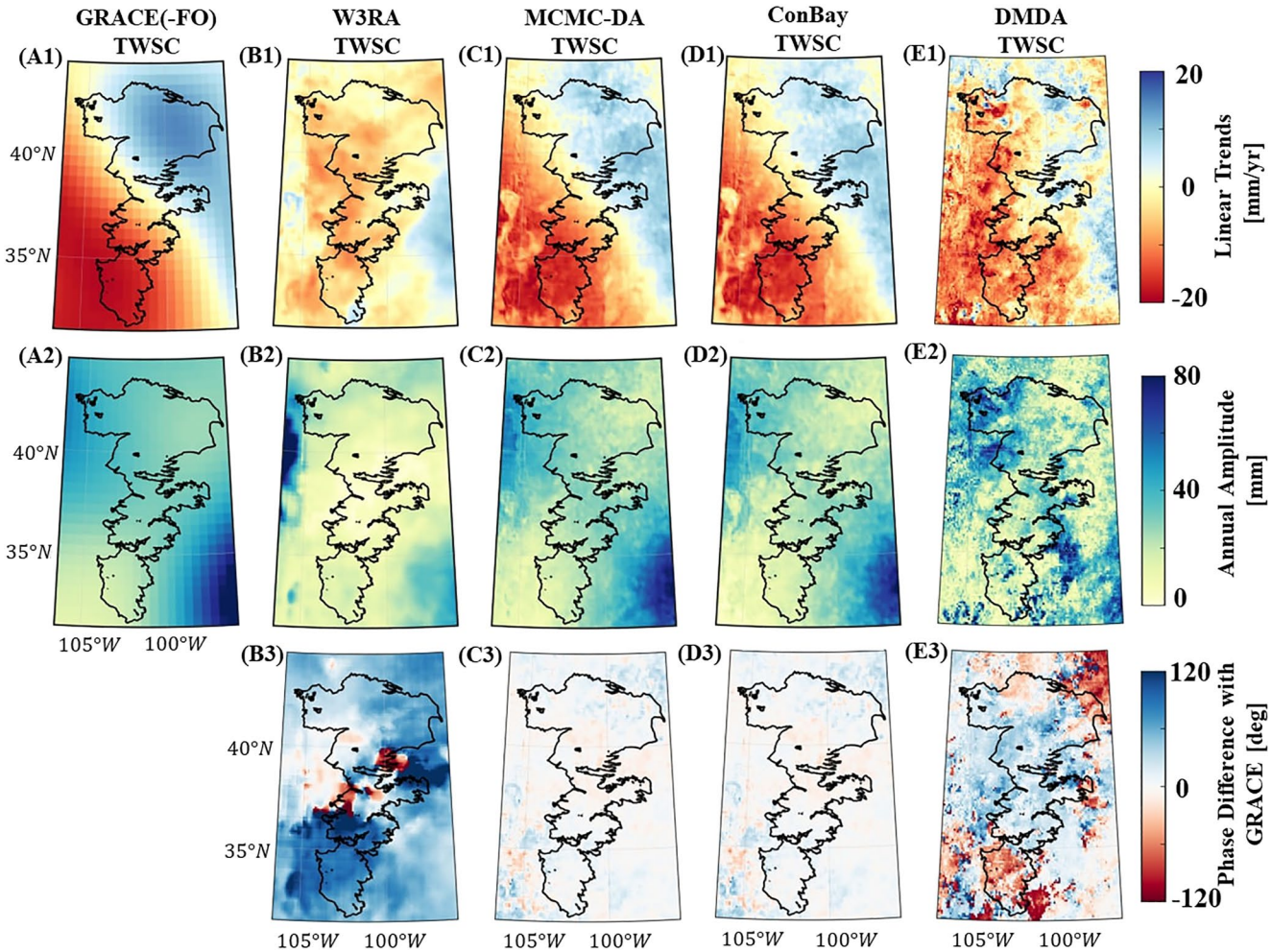

Figure 2. Long-term linear trends [mm/yr] and annual-amplitudes [mm] fitted to TWSC derived from GRACE/GRACE-FO, W3RA, ConBay, MCMC-DA, and DMDA within the High Plain (HP) aquifer, covering 2015–2021. The plots in (A1) and (A2) correspond to the GRACE/GRACE-FO observations, (B1) and (B2) correspond to the W3RA TWSC, (C1) and (C2) shows those of derived from MCMC-DA, (D1) and (D2) shows ConBay results, and (E1) and (E2) shows those of DMDA results. TWSC derived from W3RA, MCMC-DA, ConBay, and DMDA are compared in terms of phase differences with GRACE/GRACE-FO TWSC in (B3), (C3), (D3), and (E3), respectively.

between the annual component of the modeled and measured TWSC are reduced to be between ±5° after merging remote sensing observations with W3RA through the MCMC-DA and ConBay.

Long-term linear trends and annual amplitude of DMDA TWSC are shown in Figure 2 (E1) and (E2), respectively. The results show that, after implementing DMDA, the strong negative trends and the positive trends of GRACE/GRACE TWSC are successfully introduced to the W3RA TWSC in more than 74% of the HP region. The annual amplitude of model-derived TWSC is changed after implementing the DMDA, mostly in southeast and northwest of HP. However, compare to the MCMC-DA and ConBay, DMDA shows larger RMSD with GRACE/GRACE-FO TWSC, that is, ∼15 mm on average. DMDA also reduces the phase differences between modeled and measured TWSC to be between ±50° in most of the HP region. From the obtained results, it can be seen that MCMC-DA and ConBay performed better than DMDA to update W3RA TWSC with respect to the GRACE/GRACE-FO data.

4.2. Basin Average of Water Storage Changes Within the HP Aquifer

The basin averaged time series of TWSC derived from the original W3RA, DMDA, MCMC-DA, ConBay, and GRACE/GRACE-FO are presented in Figure 3 (A). From the obtained results, we can see the better performance

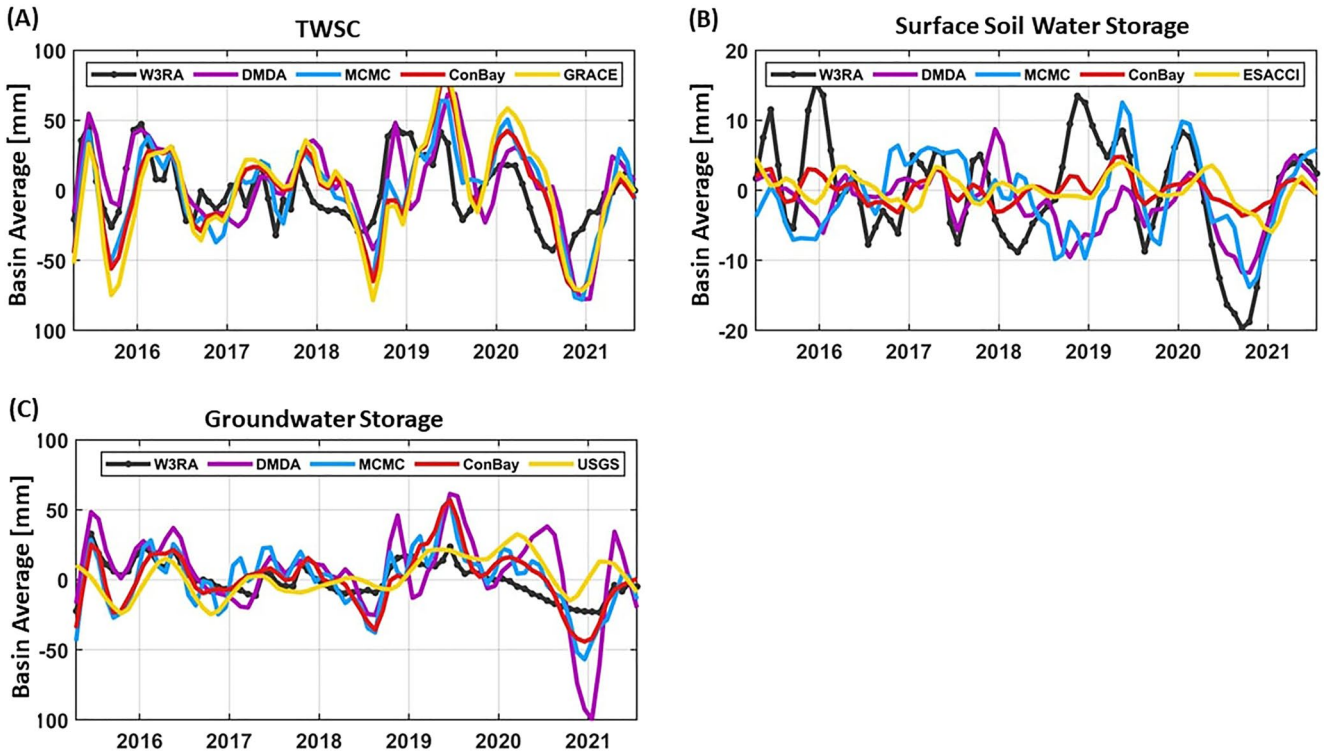

Figure 3. Basin averaged time series of (A) TWSC, (B) surface soil water, and (C) groundwater storage changes within the HP aquifer between 2015 and 2021. The basin average of water storage changes is shown for the W3RA, DMDA, MCMC-DA and ConBay. The basin averaged GRACE/GRACE-FO TWSC, ESA CCI soil water storage, and in-situ USGS groundwater storage changes are also shown in plots (A), (B), and (C), respectively.

of the MCMC-DA and ConBay, compared to DMDA. For example, considerable differences, more than 50 mm, between GRACE/GRACE-FO TWSC and DMDA TWSC can be seen in the years 2015 and 2019, while this value is close to zero for those of the MCMC-DA and ConBay approaches.

The averaged time series of surface soil water storage derived from merging remote sensing data with W3RA are also compared against those of ESA CCI product and the original model output in Figure 3 (B). Considerable changes can be seen in surface soil water storage after implementing all three merging approaches. For example, it can be seen that the strong negative magnitude (~ -20 mm) of W3RA soil water storage in years 2020–2021 is reduced to be -10 mm after implementing DMDA and MCMC-DA, and less than -4 mm after the ConBay implementation. It can be also seen that using SMAP soil data to constrain surface soil water storage considerably affects the signal separation results. For example, the RMSD between the basin averaged MCMC-DA and ESA CCI is estimated to be 7.5 mm, while the one between ConBay and ESA CCI is estimated to be less than 2.5 mm within the HP aquifer.

The basin averaged time series of groundwater storage changes of MCMC-DA and ConBay are compared against those of USGS data in Figure 3 (C). The results indicate better agreement between the groundwater storage of ConBay and the USGS data within HP, compared to those of MCMC-DA, where the temporal correlation coefficients and RMSD between the basin average of ConBay and USGS are estimated to be 0.62 and 15.6 mm, and those of between MCMC-DA and USGS are estimated to be 0.48 and 22.5 mm, respectively. It can be also seen that after implementing DMDA, the negative correlations between W3RA groundwater storage and USGS data is increased from -0.4 to 0.3. However, the large RMSD between DMDA and USGS (e.g., 80 mm in year 2021) indicates that MCMC-DA and ConBay performed better than DMDA to improve model-derived groundwater storage changes with respect to the validation data set.

In the following, we assess the performance of ConBay and MCMC-DA, in estimating individual water storage changes with the focus on surface soil water storage (top layer) and groundwater storage. Surface soil water is important to understand the land-atmosphere interactions (Brocca et al., 2017; Levine et al., 2016) and groundwater storage is the main source for agricultural and public water supplies in HP (Maupin & Barber, 2005;

Thelin & Heimes, 1987). In this study, we also provide a simple comparison between the shallow soil water storage of MCMC-DA and ConBay, which can be important for deep-rooted vegetation studies (see e.g., Gaines et al., 2016; Jochen Schenk, 2005). Long-term linear trends and annual amplitude of the DMDA surface soil water and groundwater storage changes, as well as their validations against ESA CCI soil water and in-situ USGS groundwater data are presented in Appendix A and Appendix B, respectively.

4.3. Surface Soil Water Storage Derived From Bayesian Data Assimilation

The long-term linear trends and annual amplitudes of surface soil water storage derived from W3RA, before and after the Bayesian merging (MCMC-DA and ConBay), are compared with SMAP (used as constraining data set) and the ESA CCI soil data (independent validation data set) in Figure 4. ESA CCI and SMAP soil moisture data shows decreasing surface soil water storage with the small negative trends of −1.5 mm/yr in almost 70% of the HP region. Increasing soil water storage with the positive trends of up to 1.8 mm/yr is also detected by the ESA CCI and SMAP in some parts of the east and northeast of HP (see Figure 4 (A1) and (B1), respectively). The surface soil water storage derived from W3RA, on other hand, shows strong negative trends (∼ −4 mm/yr) in almost 75% of the region (Figure 4 (C1)). It can be seen that the positive trends of remote sensing soil data in the northeast of the HP region are not reflected in the original W3RA simulations. Instead, in southeast, positive trends of up to 2 mm/yr are simulated by W3RA, which cannot be seen in the remote sensing data. The ESA CCI and SMAP remote sensing data indicate annual amplitudes of 4.2 mm for the surface coil water storage (Figure 4 (A2) and (B2), respectively). W3RA, however, simulates annual amplitudes of ∼12 mm in the south and northwest, and less than 3 mm in the east part of HP.

As shown in Figure 4 (B3) and (B4), there is a high agreement between the SMAP and ESA CCI soil water observation, where the RMSD values are estimated to be less than 4 mm, and the temporal correlation coefficients are bigger than 0.7 in almost 80% of the study region. The RMSD and correlation coefficients between W3RA and ESA CCI are also presented in Figure 4 (C3) and (C4). High correlation coefficients are visible (higher than 0.8), but significant differences are obtained in terms of RMSD (Figure 4 (C4)), with the mean value of 13 mm.

Figure 4 (D1) demonstrates that the linear trends of W3RA surface soil water are considerably changed after merging with GRACE/GRACE-FO TWSC through MCMC-DA. For example, the negative trends in the northern part of HP (∼ −4 mm/yr) are changed to positive (up to 3.2 mm), and the strong negative trends in the southern part of HP are decreased from −4 mm/yr to −2 mm/yr. Moreover, in Figure 4 (D2), we can see changes in the annual amplitudes of the soil water storage of W3RA after implementing the MCMC-DA. For example, in the northwestern and southern parts of HP, the strong annual amplitude of ∼12 mm is decreased to be less than 5 mm, which results in reducing the RMSD with the ESA CCI soil water storage, for example, from 18 to 8 mm (Figure 4 (D3)). The small RMSD values between the ESA CCI and MCMC-DA, compared to those of W3RA, confirm that merging GRACE/GRACE-FO with W3RA, improves the magnitude of surface soil water storage with respect to the validation data set. After MCMC-DA, however, the positive correlations between W3RA and ESA CCI are changed to negative in the northern HP, from for example, 0.7 to −0.1 (Figure 4 (C4) and (D4)), which might be caused by the introduced positive linear trends to the surface soil water storage by GRACE/GRACE-FO in the MCMC-DA, where no soil moisture is assimilated (shown in Figure 4 (D1)).

The ConBay merging of GRACE/GRACE-FO and SMAP with W3RA demonstrates significant improvements in both linear trends (Figure 4 (E1)) and seasonality of the surface soil data (the annual amplitude is shown in Figure 4 (E2)). In addition, the RMSD between the soil water storage of ConBay and ESA CCI is estimated to be less than 3 mm and the correlation coefficients bigger than 0.6 in almost entire HP (Figure 4 (E3) and (E4)). The small RMSD and high correlations indicate a better performance of the ConBay approach compared to MCMC-DA, where positive trends have been introduced to the upper soil layer by GRACE/GRACE-FO that are not visible in the observed (satellite) soil moisture data. We also found that the amount of updates introduced to the surface-soil water storage of W3RA decreased from 21% of the total updates to 8% of the total updates after implementing the constrain equation using SMAP data (see also Table 1).

In Table 1, the percentage of the updated values ($S_k\%$) introduced to the W3RA water storage components after implementing Bayesian merging techniques is estimated as:

$$S_k\% = \frac{\sigma^2\left(\hat{Z}_k - Z_k\right)}{\Sigma_{k=1}^{K}\sigma^2\left(\hat{Z}_k - Z_k\right)} \times 100 \tag{5}$$

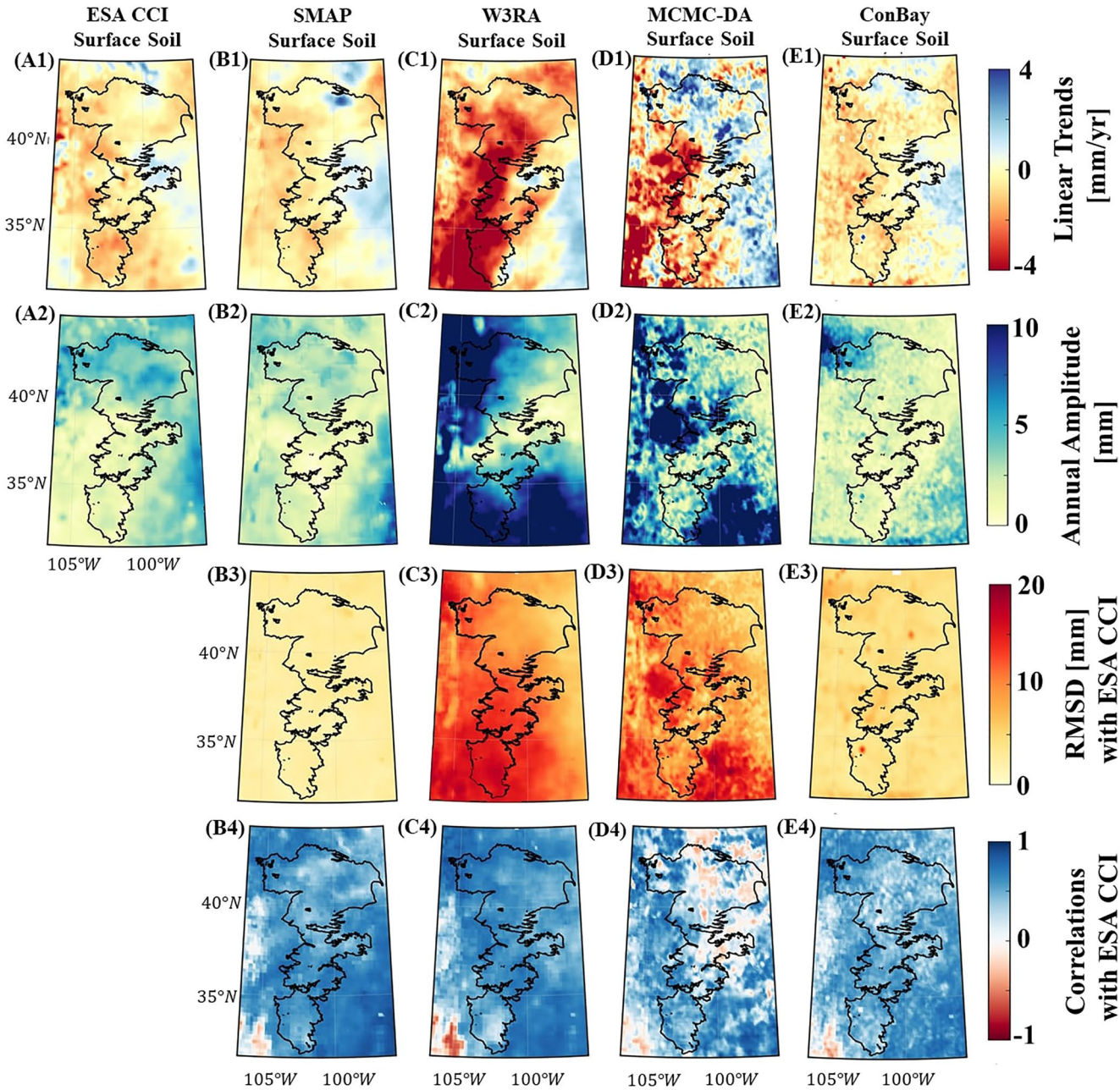

Figure 4. Long-term linear trends [mm/yr] and annual-amplitudes [mm] fitted to the surface soil water storage (top layer) derived from the ESA CCI, SMAP, W3RA, MCMC-DA and ConBay within the High Plain (HP) aquifer, covering 2015–2021. The plots in (A1) and (A2) correspond to the ESA CCI observations, (B1) and (B2) correspond to the SMAP data, (C1) and (C2) shows those of derived from the original W3RA, while (D1) and (D2) show the MCMC-DA results, and (E1) and (E2) presented ConBay results. Surface soil water storage derived from the SMAP, W3RA, MCMC-DA, and ConBay are compared in terms of RMSD with the ESA CCI in (B3), (C3), (D3), and (E3), and in terms of correlation coefficients in (B4), (C4), (D4), and (E4), respectively.

In Equation (5), σ^2 indicates the variance estimates, Z_k shows water storage components derived from W3RA, and $\hat{Z}_k$ is the Bayesian-derived water storage components, where $K = 1, 2, \ldots K$ identify the individual water storage changes such as those of surface-soil water, shallow-soil water, and groundwater storage changes.

4.4. Groundwater Storage Derived From Bayesian Data Assimilation

To assess the hypothesis that the iterative assimilation of SMAP soil moisture affects the estimation of both surface and sub-surface storage through temporal dependencies, we validate the results of MCMC-DA and

Table 1
The Percentage of Updated Values (S_k%) Introduced to the W3RA Hydrological Compartments by the MCMC-DA and ConBay Approached

Water storage compartment	MCMC-DA S_k%	ConBay S_k%
Snow + Surface water (Lakes, Rivers, …)	1%	1%
Surface-soil water	21%	8%
Shallow-soil water	21%	27%
Groundwater	57%	64%
TWSC	100%	100%

Note. Values are calculated as percentage of total variance.

ConBay against in-situ USGS groundwater observations. As can be seen in Figure 5 (A1), most of the USGS groundwater stations, with reliable observations (see Section 2.5 for more details), are located in the northern part of HP (Nebraska), which mostly show a increasing groundwater storage with linear trends up to 10 mm/yr between 2015 and 2021. W3RA, however, shows that groundwater storage is decreasing with negative trends between −2 and −8 mm/yr (Figure 5 (B1)) in almost all the HP aquifer region. After merging W3RA with remote sensing observations through MCMC-DA and ConBay, the strong negative and positive trends (∼±10 mm/yr) can be seen in Southern and Northern HP, respectively, which are similar to those of USGS in-situ observations (Figure 5 (C1) and (D1)). From the obtained results, we find that the major updates, which are introduced to the model by

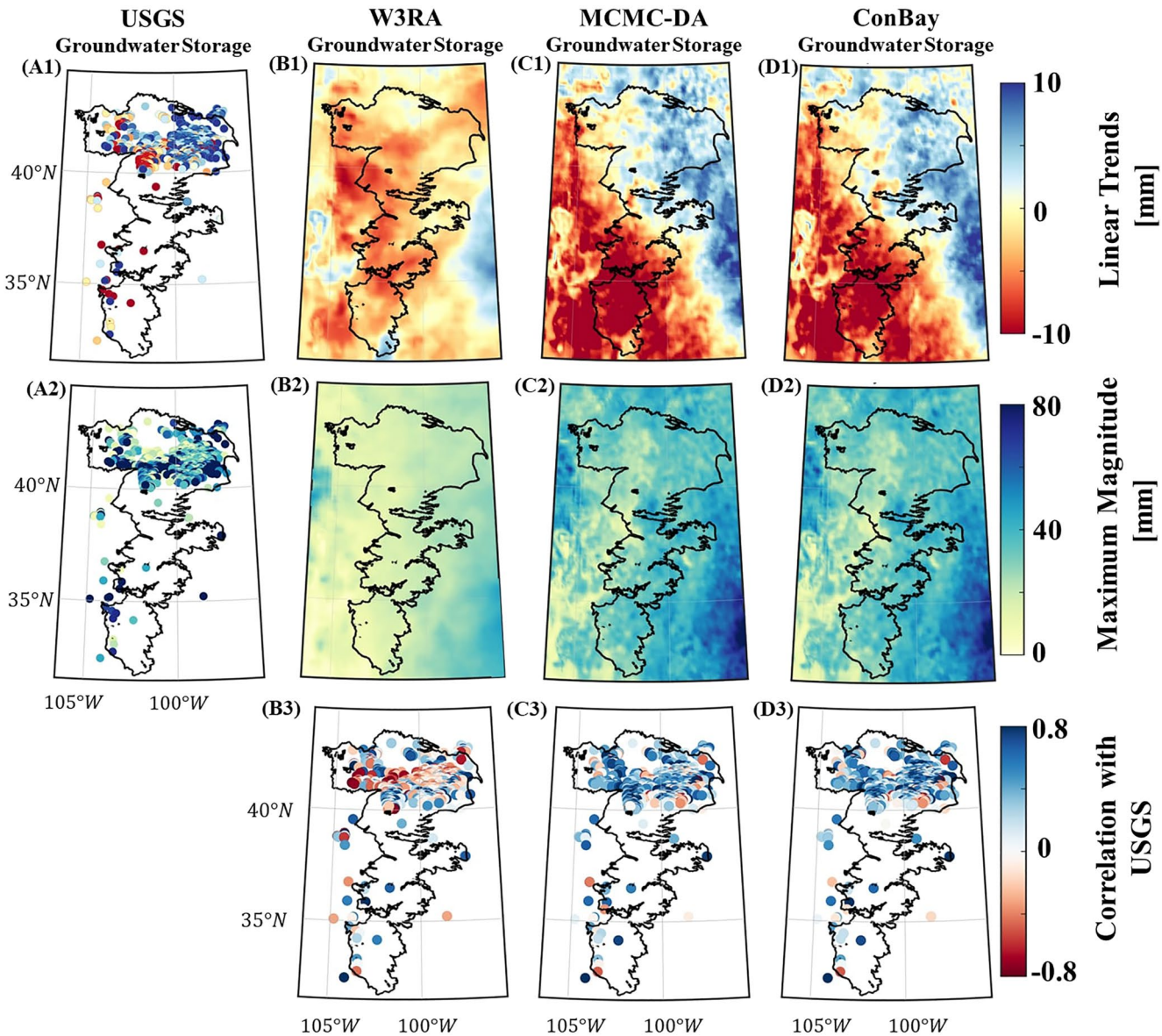

Figure 5. Long-term linear trends [mm/yr] and annual-amplitudes [mm] fitted to the groundwater storage (top layer) derived from SMAP, W3RA, MCMC-DA and ConBay within the HP aquifer, covering 2015–2021. The plots in (A1) and (A2) correspond to the USGS in-situ observations, while (B1) and (B2) correspond to the W3RA, (C1) and (C2) shows those of derived from MCMC-DA, and (D1) and (D2) shows ConBay results. Groundwater storage derived from W3RA, MCMC-DA, and ConBay are compared in terms of correlation coefficients with USGS in (B3), (C3), and (D3), respectively.

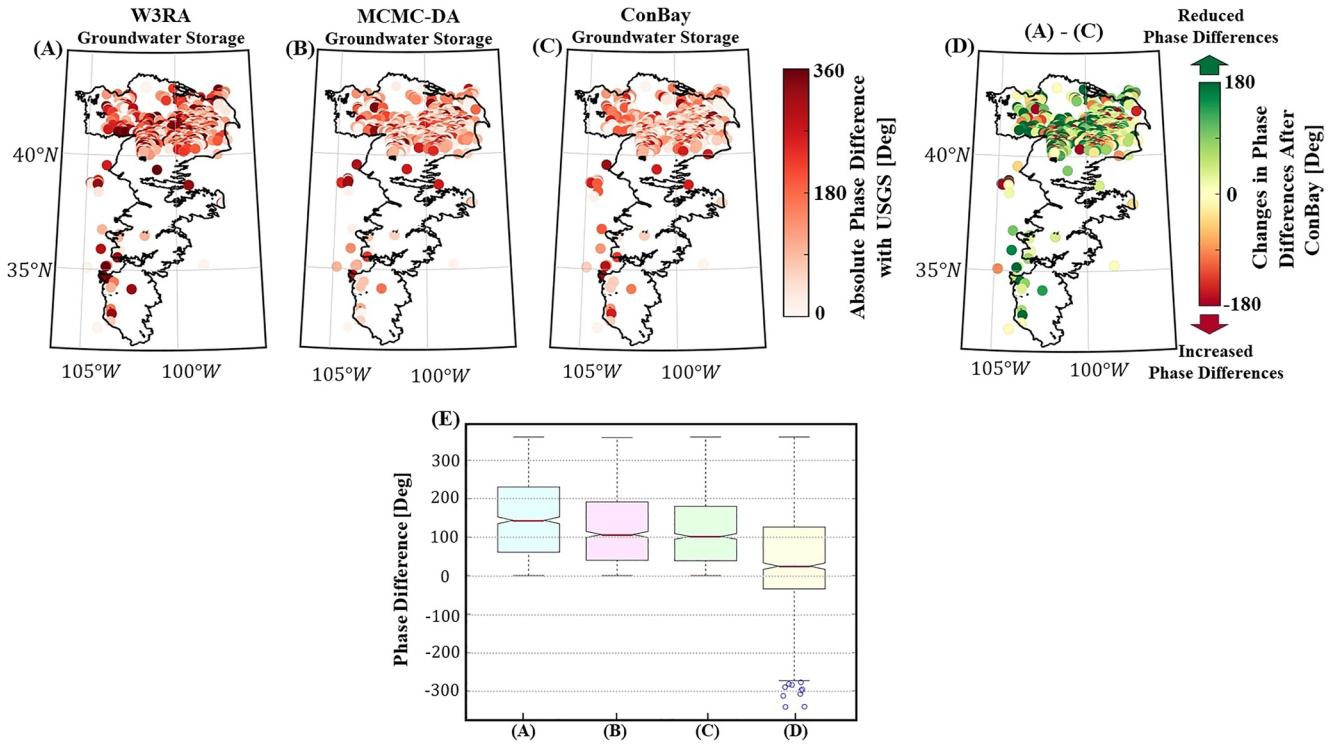

Figure 6. The absolute values of phase differences between the USGS groundwater storage and (A) the original W3RA groundwater storage, (B) the MCMC-DA groundwater, and (C) the ConBay groundwater storage changes. Differential values between (A) and (C) are shown in (D), where the positive values indicate that (A) is bigger than (C). The statistical comparison of the plots in (A), (B), (C), and (D) are presented in (E). The tops and bottoms of each box indicate the 25th and 75th percentiles (the first and third quartiles), and the red lines show the median values of each group.

GRACE/GRACE-FO TWSC, are assigned to the groundwater storage changes (see Table 1), in which the magnitudes of changes in groundwater storage (Figure 5 (B2)) is increased from 21 to 53 mm after merging W3RA with remote sensing data through MCMC-DA and ConBay (Figure 5 (C2) and (D2)). This is caused by the fact that changes in the groundwater storage of W3RA forms the biggest portion of TWSC (43% of total variance) within HP. Therefore, groundwater storage gains the highest value of temporal weights trough the MCMC-DA and ConBay, which leads to the highest value of updates from GRACE/GRACE-FO TWSC.

Bayesian-derived groundwater storage changes are found to be closer to the USGS in-situ observations with the mean annual amplitude of ~52 mm. The original W3RA shows negative correlations (~ -0.4) with USGS data, in 68% of the validation points (Figure 5 (B3)), while MCMC-DA and ConBay show positive correlations (higher than 0.5) with USGS in 88% of the validation points (Figure 5 (C3) and (D3)). Therefore, it can be stated that the merging, through MCMC-DA and ConBay, introduces useful updates of groundwater storage changes in the model. However, the amount of updates introduced to the groundwater storage between MCMC-DA (57% of total variance) and ConBay (64% of total variance) is slightly different (see Table 1). This also causes small differences between their linear trends (±2 mm/yr) and seasonality (~3.5 mm) and, therefore, it is resulted in the smaller RMSD and higher correlation coefficients between the groundwater storage of ConBay and USGS (see also the results in Section 4.2).

To demonstrate how the ConBay approach improves the phase shift (timing) of the groundwater storage, the absolute values of phase differences between annual amplitude of in situ (USGS) and model (W3RA) groundwater storage are shown in Figure 6 (A), and the results are compared with those of between the USGS and MCMC-DA groundwater storage in Figure 6 (B), and between the USGS and ConBay groundwater storage in Figure 6 (C). The difference between Figures 6 (A) and (C) is presented in Figure 6 (D), where the positive values show that the phase amplitude of groundwater storage improves after implementing the ConBay. In general, after MCMC-DA and ConBay, the phase differences between modeled and observed groundwater storage changes are decreased for 68% of the validation points (see Figure 6 (E) yellow box). This improvement can also be seen

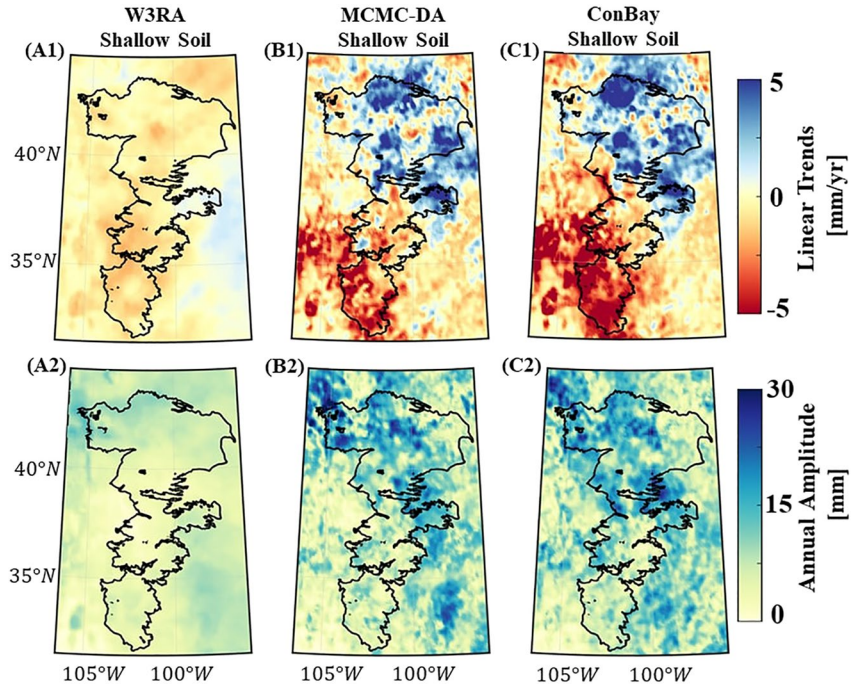

Figure 7. Long-term linear trends [mm/yr] and annual-amplitudes [mm] fitted to the shallow soil water storage derived from the original W3RA, MCMC-DA, and ConBay within the HP aquifer, covering 2015–2021. The plots in (A1) and (A2) correspond to W3RA, while (B1) and (B2) represent MCMC-DA, and (C1) and (C2) show those of ConBay.

in Figure 6 (E), where the median of the absolute phase differences (red lines) between the annual amplitude of W3RA and USGS groundwater data (blue box) is reduced from 150° to 100° after implementing the ConBay approach (green box).

4.5. Shallow Soil Water Storage Derived From Bayesian Data Assimilation

In Table 1, it can be seen that after groundwater storage changes, shallow soil water (5 − 100 cm) gains the highest values of updates from GRACE/GRACE-FO TWSC through the ConBay approach. The obtained results indicate that in the MCMC-DA, where we do not use SMAP data, 21% of the total updates is assigned to the surface soil water storage. This value is reduced to 8% after constraining the surface soil water storage using SMAP. TWSC derived from the MCMC-DA and ConBay (as a summation of updated water storage compartments) are almost equal and are fitted to the GRACE/GRACE-FO TWSC (with the RMSD of less than 5 mm). Therefore, the total updates introduced to the model should be equal for both techniques. Therefore, the reduction in the updated values of surface soil water storage results in increasing the updated values of other compartments, such as groundwater storage and shallow soil water storage. This shows the importance of acknowledging the connection between different storage compartments.

Long term linear trends and annual amplitudes fitted to the shallow soil water before and after assimilation are presented in Figure 7. Due to the lack of available in-situ or remote sensing observations to validate shallow soil water, here we only focus on a comparison between the shallow soil of the MCMC-DA and ConBay compared with the original W3RA.

The original W3RA shows a decreasing trend in the shallow soil water of HP at −2.3 mm/yr, on average, during 2015–2021 (Figure 7 (A1)). Merging GRACE/GRACE-FO and SMAP with the model, through MCMC-DA and ConBay (Figure 7 (B1) and (C1), respectively), significantly change these trends to the strong positive trends in the northern part of HP (up to 5 mm/yr) and introduces strong negative trends to the southern HP (down to −5 mm/yr). Other differences can be seen in the annual amplitude of the shallow soil water storage before and after assimilation, where the annual amplitude of ∼7 mm in W3RA (Figure 7 (A2)) is increased to the values between 15 and 25 mm, for both MCMC-DA and ConBay (Figure 7 (B2) and (C2)).

Although similar spatial patterns can be seen in the linear trend and the annual amplitude of MCMC-DA and ConBay, there are some differences between them. For example, a comparison between Figure 7 (B1) and (C1)) demonstrates that the shallow soil water of ConBay contains stronger positive and negative linear trends in the north and south parts of HP compared to those of MCMC-DA. The annual amplitude of the shallow soil water of ConBay is also bigger than those of MCMC-DA. These differences capture 21% and 27% of the total variance of the difference between updated values from MCMC-DA and ConBay, respectively, see Table 1.

5. Summary and Conclusion

In this study, the performance of our newly developed "Constrained Bayesian (ConBay)" optimization approach (Forootan & Mehrnegar, 2022) was assessed to jointly merge the GRACE/GRACE-FO TWSC and the SMAP soil moisture data with the output of a large-scale hydrological model. To evaluate the contribution of the SMAP data on the final water storage estimates an 'unconstrained' Bayesian merger, known as MCMC-DA, was implemented to merge only GRACE/GRACE-FO TWSC with W3RA, where the SMAP soil moisture data was not used to constrain the estimation of the surface soil water storage. This allowed to evaluate the contribution of SMAP data on the final water storage estimates. The Dynamic Model Data Averaging (DMDA) was also implemented here to merge GRACE/GRACE-FO TWSC and SMAP soil moisture data with W3RA model outputs. This is a close implementation to that of EnKF but in the offline mode, where for DMDA it is not required to run the model after each data assimilation (merging) step. Thus, the DMDA implementation is close to that of ConBay and provides an opportunity of a fair comparison. As a case study, the High Plain (HP) aquifer in United States (US) and the period of 2015–2021 were chosen to merge the GRACE/GRACE-FO TWSC and the SMAP soil moisture data into the W3RA water balance model with the spatial resolution of $0.1° \times 0.1°$. Validations were conducted with the in-situ USGS groundwater storage observations and the European Space Agency (ESA)'s Climate Change Initiative (CCI) soil moisture products.

From the obtained results, we found that the ConBay assimilation of remote sensing observations into the hydrological model successfully updated the seasonal and inter-annual components of hydrological signals and reduced biases and phase differences between the modeled and measured TWSC, as well as individual storage compartments. We also found that using SMAP data to constrain soil water storage considerably reduced the updated values introduced to this compartment (8% of the total variance), compared to the assimilation of GRACE/GRACE-FO TWSC only in the MCMC-DA approach (21% of the total variance). Thus, the surface soil water ConBay showed a better agreement, in terms of both linear trends and seasonality, with the ESA CCI estimates compared to those derived from MCMC-DA. This is because, in MCMC-DA, the amount of updates introduced to the water storage compartment of W3RA is weighted based on the total variance of each compartment. In ConBay, however, these weights are optimized by considering the limited boundary defined by the constrain equation that uses SMAP as input. Our results confirm the finding by, for example, Tian et al. (2017); Girotto et al. (2019); Tangdamrongsub et al. (2020), who discussed the introduced unwanted signals to the surface soil water storage after implementing a GRACE-only data assimilation.

It was also found that the magnitude of updates introduced to the groundwater storage of MCMC-DA (57% of total variance) and ConBay (64% of total variance) was slightly different (Table 1), which caused the better agreement between the groundwater storage of ConBay and USGS compared to those derived from the MCMC-DA approach.

A comparison between the shallow soil water derived from MCMC-DA and ConBay demonstrated that the updated values introduced to this compartment was also different for MCMC-DA and ConBay (21% and 27% of total updates, respectively), which resulted in differences between their linear trends and seasonality. These findings are the evidence for the hypothesis that using SMAP soil moisture to constrain the updated values of surface soil water storage affects the estimation of all other hydrological compartments.

In this study, we found that the fully Bayesian integration of remote sensing data with the model outputs performed better than introding them individually as it is implemented in DMDA. The implementation of the DMDA, however, is more computationally efficient compared to the Bayesian DA techniques, which makes it a unique approach to integrate multiple hydrological models with GRACE/GRACE-FO data over large spatial domains and for applications with high spatial resolution. In what follows, some of the main remarks of this study are summarized.

- The RMSD between modeled TWSC and GRACE/GRACE-FO TWSC reduced considerably by 73% (from 38 to 5 mm on average), on average after implementing MCMC-DA and ConBay.
- The phase differences between the annual amplitude of GRACE/GRACE-FO TWSC and modeled TWSC were reduced from ±120° to less than ±5° after implementing both MCMC-DA and ConBay.
- The surface soil water storage of ConBay showed smaller RMSD (less than 4 mm) and higher correlation coefficients (bigger than 80%) with the ESA CCI compared to those of MCMC-DA.
- The negative correlations (−0.4) between the USGS observations and modeled groundwater storage were significantly changed to the positive values (~0.6) after merging them with the GRACE/GRACE-FO TWSC and the SMAP soil data.
- The RMSD between the groundwater of ConBay and that of USGS was estimated to be 15.6 mm, while the one between MCMC-DA and USGS data was estimated to be 22.5 mm.

Appendix A: Evaluation of DMDA Surface Soil Water Storage Changes

The long-term linear trends and annual amplitudes of the surface soil water storage derived from DMDA are shown in Figures A1 (A) and (B), respectively. Same as MCMC-DA and ConBay, DMDA surface soil water storage is validated against the ESA CCI soil data in terms of temporal correlation coefficients and RMSD in Figures A1 (C) and (D), respectively. The obtained results indicate that after implementing DMDA to merge GRACE/GRACE-FO TWSC and SMAP soil data with W3RA, the strong negative trends of W3RA surface soil water storage (Figure 4 (C1)) is reduced by ~50% (from −4 to −2 mm/yr) in southern HP, which is closer to those of the ESA CCI and SMAP data (Figure 4 (A1) and (B1)). It can also be seen that the small positive trends of SMAP in the northern part of HP is reflected in the DMDA results. The annual amplitude of the DMDA surface soil water storage is closer to those of SMAP and ESA CCI compared to the original model outputs (see Figure 4 (A2), (B2), (C2)). For example, in the northwestern and southern parts of HP, the strong annual amplitude of W3RA (12 mm) is decreased to be less than 8 mm, which results in reducing RMSD with the ESA CCI soil water storage (compare Figures A1 (C) and 4 (C3)). After DMDA, however, the positive correlations between the W3RA and ESA CCI are negatively changed in ~50% of HP (Figure A1 (D)). It might be caused by the weakness of the DMDA to estimate temporal dependency between the water storage compartments, which negatively affects the estimations of all individual storage changes.

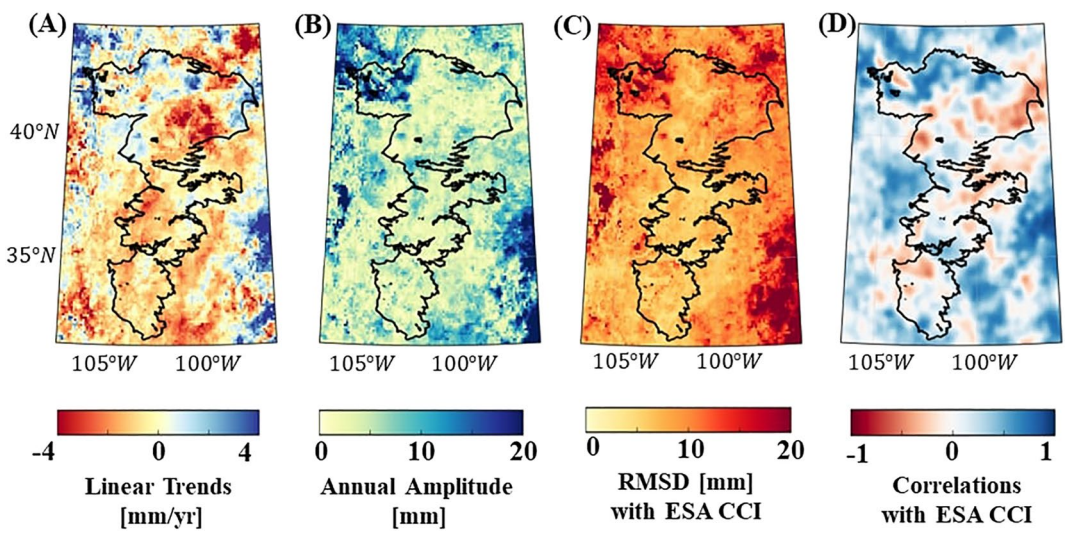

Figure A1. Long-term linear trends [mm/yr] and annual-amplitudes [mm] fitted to the surface soil water storage derived from DMDA within the HP aquifer, covering 2015–2021 (plots (A) and (B)). Surface soil water storage derived from DMDA is compared with the ESA CCI soil data in terms of RMSD and correlation coefficients in (C) and (D), respectively.

Appendix B: Evaluation of DMDA Groundwater Storage Changes

Long-term linear trends and maximum magnitude of the DMDA groundwater storage changes, presented in Figures B1 (A) and (B), indicate that after implementing DMDA the strong negative and positive trends of GRACE/GRACE-FO TWSC in the southwest and northeast of HP are introduced to the groundwater storage changes. The maximum magnitude of the DMDA groundwater storage, however, was not changed considerably after DMDA. The strong magnitude in the west of HP can be interpreted as uncertainties, since it does not appear in the observation data. The negative correlation coefficients between the in-situ USGS and groundwater storage changes of W3RA (Figure 5 (B3)) are increased up to 0.3 after implementing DMDA, which are smaller than those of MCMC-DA and ConBay shown in Figure 5. Negative correlations between the USGS and the groundwater storage of DMDA in the northwest of HP can be related to the positive trends of USGS in this region (Figure 5 (A1)), while DMDA shows strong negative trends (Figure B1 (A)). MCMC-DA and ConBay, however, show positive correlation coefficients, up to 0.8, in most parts of HP, which indicates that the fully Bayesian integration of remote sensing data with model works more efficient compare to DMDA.

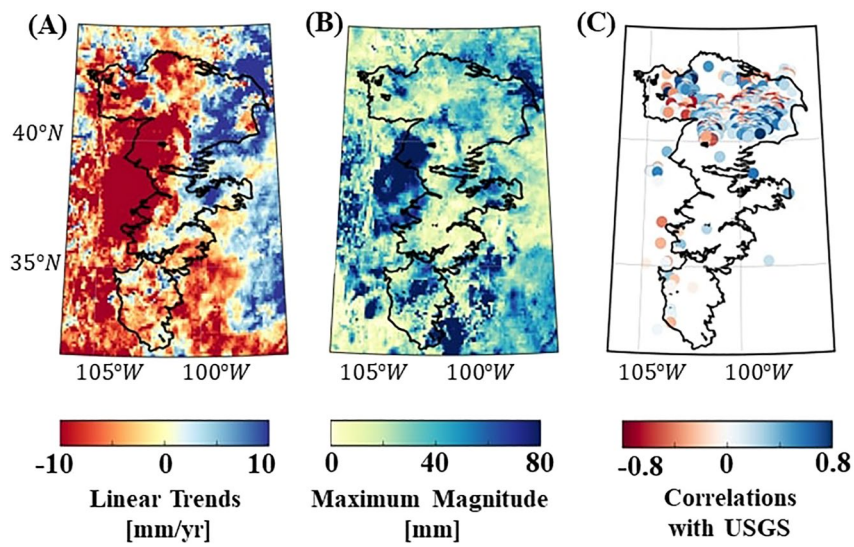

Figure B1. Long-term linear trends [mm/yr] and annual-amplitudes [mm] fitted to the groundwater storage derived from DMDA within the HP aquifer, covering 2015–2021 (plots (A) and (B), respectively). The groundwater storage derived from DMDA is compared with the in-situ USGS groundwater storage in terms of correlation coefficients in (C).

Data Availability Statement

All the data supporting the findings of this study are available within the article. Reconstructed GRACE/GRACE-FO TWSC within the HP aquifer can be downloaded from https://doi.org/10.6084/m9.figshare.23735724.v1. The monthly average of enhanced L3 soil moisture product (O'Neill et al., 2021), version 5, retrieved by the Soil Moisture Active Passive (SMAP) radiometer, is used in this study to constrain the estimation of near-surface soil moisture (top soil layer). For this study, the original W3RA code is modified to simulate TWSC compartment within the HP aquifer using the ERA5 forcing data. The original codes can be downloaded from http://wald.anu.edu.au/challenges/water/w3-and-ozwald-hydrology-models/. The outputs of W3RA within the HP aquifer can be downloaded from https://doi.org/10.6084/m9.figshare.23735724.v1. The version 07.1 of the daily ESA CCI soil moisture with a spatial resolution of 0.25° × 0.25° and covering the period of 1978–2021 is downloaded from the ESA website http://www.esa-soilmoisture-cci.org as an independent validation data set for this study. The groundwater level data in the HP aquifer, as an independent validation data set, are collected from the US Geological Survey (USGS) groundwater network https://water.usgs.gov/ogw/networks.html. The code for implementing the ConBay approach to merge multi-remote sensing EO data with model outputs is written in the MATLAB based on the formulations in Forootan and Mehrnegar (2022). The code for MCMC-DA is also written in MATLAB based on the formulations in Mehrnegar, Jones, Singer, Schumacher, Jagdhuber, et al. (2020).

TWSC compartments derived from ConBay and MCMC-DA approaches can be downloaded from https://doi.org/10.6084/m9.figshare.23735724.v1.

Acknowledgments
This work was supported by the Danmarks Frie Forskningsfond [10.46540/2035-00247B]. N. Mehrnegar is grateful to the supports by the Marie Skłodowska-Curie Actions (MSCA) postdoctoral Fellowship [101068561]. We would also like to express our gratitude to Professor Albert Van Dijk for providing the original W3RA water balance codes.

References

Andreadis, K. M., & Lettenmaier, D. P. (2006). Assimilating remotely sensed snow observations into a macroscale hydrology model. *Advances in Water Resources*, *29*(6), 872–886. https://doi.org/10.1016/j.advwatres.2005.08.004

Argus, D. F., Peltier, W., Drummond, R., & Moore, A. W. (2014). The Antarctica component of postglacial rebound model ice-6g_c (vm5a) based on gps positioning, exposure age dating of ice thicknesses, and relative sea level histories. *Geophysical Journal International*, *198*(1), 537–563. https://doi.org/10.1093/GJI/GGU140

Awange, J. L., Ferreira, V., Forootan, E., Khandu, S. A., Agutu, N., & He, X. (2016). Uncertainties in remotely sensed precipitation data over Africa. *International Journal of Climatology*, *36*(1), 303–323. https://doi.org/10.1002/joc.4346

Bain, A., & Crisan, D. (2008). *Fundamentals of stochastic filtering*. Vol. 60. Springer Science & Business Media.

Bernstein, D. S. (2005). *Matrix mathematics: Theory, facts, and formulas with application to linear systems theory (Vol. 41)*. Princeton University Press.

Bezděk, A., Sebera, J., Teixeira da Encarnação, J., & Klokočník, J. (2016). Time-variable gravity fields derived from GPS tracking of Swarm. *Geophysical Journal International*, *205*(3), 1665–1669. https://doi.org/10.1093/gji/ggw094

Blankenship, C. B., Case, J. L., Zavodsky, B. T., & Crosson, W. L. (2016). Assimilation of SMOS retrievals in the land information system. *IEEE Transactions on Geoscience and Remote Sensing*, *54*(11), 6320–6332. https://doi.org/10.1109/TGRS.2016.2579604

Brocca, L., Ciabatta, L., Massari, C., Camici, S., & Tarpanelli, A. (2017). Soil moisture for hydrological applications: Open questions and new opportunities. *Water*, *9*(2), 140. https://doi.org/10.3390/w9020140

Bruckler, L., Witono, H., & Stengel, P. (1988). Near surface soil moisture estimation from microwave measurements. *Remote Sensing of Environment*, *26*(2), 101–121. https://doi.org/10.1016/0034-4257(88)90091-0

Butler, J., Bohling, G., Whittemore, D., & Wilson, B. (2020). A roadblock on the path to aquifer sustainability: Underestimating the impact of pumping reductions. *Environmental Research Letters*, *15*(1), 014003. https://doi.org/10.1088/1748-9326/ab6002

Cano, A., Núñez, A., Acosta-Martinez, V., Schipanski, M., Ghimire, R., Rice, C., & West, C. (2018). Current knowledge and future research directions to link soil health and water conservation in the ogallala aquifer region. *Geoderma*, *328*, 109–118. https://doi.org/10.1016/j.geoderma.2018.04.027

Castellazzi, P., Martel, R., Rivera, A., Huang, J., Pavlic, G., Calderhead, A. I., et al. (2016). Groundwater depletion in Central Mexico: Use of GRACE and InSAR to support water resources management. *Water Resources Research*, *52*(8), 5985–6003. https://doi.org/10.1002/2015WR018211

Chan, S. K., Bindlish, R., O'Neill, P. E., Njoku, E., Jackson, T., Colliander, A., et al. (2016). Assessment of the SMAP passive soil moisture product. *IEEE Transactions on Geoscience and Remote Sensing*, *54*(8), 4994–5007. https://doi.org/10.1109/TGRS.2016.2561938

Chen, J., Cazenave, A., Dahle, C., Llovel, W., Panet, I., Pfeffer, J., & Moreira, L. (2022). Applications and challenges of grace and grace follow-on satellite gravimetry. *Surveys in Geophysics*, *43*, 1–345. https://doi.org/10.1007/s10712-021-09685-x

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), 237–245. https://doi.org/10.1007/s00190-006-0104-2

Chib, S., & Greenberg, E. (1995). Understanding the metropolis-hastings algorithm. *The American Statistician*, *49*(4), 327–335. https://doi.org/10.1080/00031305.1995.10476177

Das, N. N., Entekhabi, D., Dunbar, R. S., Chaubell, M. J., Colliander, A., Yueh, S., et al. (2019). The smap and copernicus sentinel 1a/b microwave active-passive high resolution surface soil moisture product. *Remote Sensing of Environment*, *233*, 111380. https://doi.org/10.1016/j.rse.2019.111380

Dorigo, W., Wagner, W., Albergel, C., Albrecht, F., Balsamo, G., Brocca, L., et al. (2017). ESA CCI soil moisture for improved Earth system understanding: State-of-the art and future directions. *Remote Sensing of Environment*, *203*, 185–215. https://doi.org/10.1016/j.rse.2017.07.001

Du, Y., Ulaby, F. T., & Dobson, M. C. (2000). Sensitivity to soil moisture by active and passive microwave sensors. *IEEE Transactions on Geoscience and Remote Sensing*, *38*(1), 105–114. https://doi.org/10.1109/36.823905

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*(6), 1285–1309. https://doi.org/10.1007/s10712-014-9309-8

Entekhabi, D., Njoku, E. G., O'Neill, P. E., Kellogg, K. H., Crow, W. T., Edelstein, W. N., et al. (2010). The soil moisture active passive (smap) mission. *Proceedings of the IEEE*, *98*(5), 704–716. https://doi.org/10.1109/JPROC.2010.2043918

Evensen, G., & Van Leeuwen, P. J. (2000). An ensemble kalman smoother for nonlinear dynamics. *Monthly Weather Review*, *128*(6), 1852–1867. https://doi.org/10.1175/1520-0493(2000)128<1852:aeksfn>2.0.co;2

Fang, B., Lakshmi, V., Bindlish, R., & Jackson, T. J. (2018). Downscaling of smap soil moisture using land surface temperature and vegetation data. *Vadose Zone Journal*, *17*(1), 1–15. https://doi.org/10.2136/vzj2017.11.0198

Fang, B., Lakshmi, V., Bindlish, R., Jackson, T. J., Cosh, M., & Basara, J. (2013). Passive microwave soil moisture downscaling using vegetation index and skin surface temperature. *Vadose Zone Journal*, *12*(3), 1–19. https://doi.org/10.2136/vzj2013.05.0089

Fang, B., Lakshmi, V., Bindlish, R., Jackson, T. J., & Liu, P.-W. (2020). Evaluation and validation of a high spatial resolution satellite soil moisture product over the continental United States. *Journal of Hydrology*, *588*, 125043. https://doi.org/10.1016/j.jhydrol.2020.125043

Fay, M. P., & Proschan, M. A. (2010). Wilcoxon-Mann-Whitney or t-test? On assumptions for hypothesis tests and multiple interpretations of decision rules. *Statistics Surveys*, *4*, 1. https://doi.org/10.1214/09-SS051

Ferreira, V. G., Montecino, H. D. C., Yakubu, C. I., & Heck, B. (2016). Uncertainties of the gravity recovery and climate experiment time-variable gravity-field solutions based on three-cornered hat method. *Journal of Applied Remote Sensing*, *10*(1), 1–20. https://doi.org/10.1117/1.JRS.10.015015

Flechtner, F., Neumayer, K.-H., Dahle, C., Dobslaw, H., Fagiolini, E., Raimondo, J.-C., & Güntner, A. (2016). What can be expected from the GRACE-FO laser ranging interferometer for Earth science applications? *Surveys in Geophysics*, *33*(2), 453–470. https://doi.org/10.1007/s10712-015-9338-y

Forootan, E., Awange, J. L., Kusche, J., Heck, B., & Eicker, A. (2012). Independent patterns of water mass anomalies over Australia from satellite data and models. *Remote Sensing of Environment*, *124*, 427–443. https://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 2003–2011. *Journal of Geodesy*, *88*(5), 503–514. https://doi.org/10.1007/s00190-014-0696-x

Forootan, E., Khaki, M., Schumacher, M., Wulfmeyer, V., Mehrnegar, N., Van Dijk, A., et al. (2019). Understanding the global hydrological droughts of 2003–2016 and their relationships with teleconnections. *Science of the Total Environment*, *650*, 2587–2604. https://doi.org/10.1016/j.scitotenv.2018.09.231

Forootan, E., & Kusche, J. (2013). Separation of deterministic signals using independent component analysis (ICA). *Studia Geophysica et Geodaetica*, *57*(1), 17–26. https://doi.org/10.1007/s11200-012-0718-1

Forootan, E., & Mehrnegar, N. (2022). A hierarchical constrained bayesian (conbay) approach to jointly estimate water storage and post-glacial rebound from grace (-fo) and gnss data. *All Earth*, *34*(1), 120–146. https://doi.org/10.1080/27669645.2022.2097768

Forootan, E., Rietbroek, R., Kusche, J., Sharifi, M. A., Awange, J. L., Schmidt, M., et al. (2014). Separation of large scale water storage patterns over Iran using GRACE, altimetry and hydrological data. *Remote Sensing of Environment*, *140*, 580–595. https://doi.org/10.1016/j.rse.2013.09.025

Forootan, E., Safari, A., Mostafaie, A., Schumacher, M., Delavar, M., & Awange, J. L. (2017). Large-scale total water storage and water flux changes over the arid and semiarid parts of the Middle East from GRACE and reanalysis products. *Surveys in Geophysics*, *38*(3), 591–615. https://doi.org/10.1007/s10712-016-9403-1

Forootan, E., Schumacher, M., Mehrnegar, N., Bezděk, A., Talpe, M. J., Farzaneh, S., et al. (2020). An iterative ICA-based reconstruction method to produce consistent time-variable total water storage fields using GRACE and Swarm satellite data. *Remote Sensing*, *12*(10), 1639. https://doi.org/10.3390/rs12101639

Gaines, K. P., Stanley, J. W., Meinzer, F. C., McCulloh, K. A., Woodruff, D. R., Chen, W., et al. (2016). Reliance on shallow soil water in a mixed-hardwood forest in central Pennsylvania. *Tree Physiology*, *36*(4), 444–458. https://doi.org/10.1093/treephys/tpv113

Gelfand, A. E., & Smith, A. F. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association*, *85*(410), 398–409. https://doi.org/10.1080/01621459.1990.10476213

Geyer, C. J. (1991). *Markov chain Monte Carlo maximum likelihood*.

Gillies, R. R., & Carlson, T. N. (1995). Thermal remote sensing of surface soil water content with partial vegetation cover for incorporation into climate models. *Journal of Applied Meteorology and Climatology*, *34*(4), 745–756. https://doi.org/10.1175/1520-0450(1995)034<0745:trsoss>2.0.co;2;2

Girotto, M., De Lannoy, G. J., Reichle, R. H., & Rodell, M. (2016). Assimilation of gridded terrestrial water storage observations from GRACE into a land surface model. *Water Resources Research*, *52*(5), 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*(9), 4107–4115. https://doi.org/10.1002/2017GL072994

Girotto, M., Reichle, R. H., Rodell, M., Liu, Q., Mahanama, S., & De Lannoy, G. J. (2019). Multi-sensor assimilation of SMOS brightness temperature and GRACE terrestrial water storage observations for soil moisture and shallow groundwater estimation. *Remote Sensing of Environment*, *227*, 12–27. https://doi.org/10.1016/j.rse.2019.04.001

Gruber, A., Scanlon, T., Van der Schalie, R., Wagner, W., & Dorigo, W. (2019). Evolution of the ESA CCI Soil Moisture climate data records and their underlying merging methodology. *Earth System Science Data*, *11*(2), 1–739. https://doi.org/10.5194/essd-11-717-2019

Gutentag, E. D., Heimes, F. J., Krothe, N. C., Luckey, R. R., & Weeks, J. B. (2014). *Geohydrology of the high plains aquifer in parts of Colorado, Kansas, Nebraska, New Mexico, Oklahoma, South Dakota, Texas, and Wyoming*, *1400-B*, 63. https://doi.org/10.3133/pp1400B

Hersbach, H., & Dee, D. (2016). Era5 reanalysis is in production. *ECMWF newsletter*, *147*(7), 5–6. https://doi.org/10.24381/cds.adbb2d47

Hoeting, J. A., Madigan, D., Raftery, A. E., & Volinsky, C. T. (1999). Bayesian model averaging: A tutorial. *Statistical Science*, *14*(4), 382–417. https://doi.org/10.1214/ss/1009212519

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*(7). https://doi.org/10.1029/2011WR011291

Jagdhuber, T., Baur, M., Akbar, R., Das, N. N., Link, M., He, L., & Entekhabi, D. (2019). Estimation of active-passive microwave covariation using smap and sentinel-1 data. *Remote Sensing of Environment*, *225*, 458–468. https://doi.org/10.1016/j.rse.2019.03.021

Jochen Schenk, H. (2005). *Vertical vegetation structure below ground: Scaling from root to globe*. In *Progress in botany* (pp. 341–373). Springer. https://doi.org/10.1007/BF02257569

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, *82*(1), 35–45. https://doi.org/10.1115/1.3662552

Kerr, Y. H., Waldteufel, P., Richaume, P., Wigneron, J. P., Ferrazzoli, P., Mahmoodi, A., et al. (2012). The SMOS soil moisture retrieval algorithm. *IEEE Transactions on Geoscience and Remote Sensing*, *50*(5), 1384–1403. https://doi.org/10.1109/TGRS.2012.2184548

Kerr, Y. H., Waldteufel, P., Wigneron, J.-P., Delwart, S., Cabot, F., Boutin, J., et al. (2010). The smos mission: New tool for monitoring key elements ofthe global water cycle. *Proceedings of the IEEE*, *98*(5), 666–687. https://doi.org/10.1109/JPROC.2010.2043032

Khaki, M., Hoteit, I., Kuhn, M., Awange, J., Forootan, E., Van Dijk, A. I., et al. (2017). Assessing sequential data assimilation techniques for integrating GRACE data into a hydrological model. *Advances in Water Resources*, *107*, 301–316. https://doi.org/10.1016/j.advwatres.2017.07.001

Kitagawa, G. (1987). Non-Gaussian state? Space modeling of nonstationary time series. *Journal of the American Statistical Association*, *82*(400), 1032–1041. https://doi.org/10.1080/01621459.1987.10478534

Koller, D., & Friedman, N. (2009). *Probabilistic graphical models: Principles and techniques*. MIT Press.

Kusche, J., Schmidt, R., Petrovic, S., & Rietbroek, R. (2009). Decorrelated GRACE time-variable gravity solutions by GFZ, and their validation using a hydrological model. *Journal of Geodesy*, *83*(10), 903–913. https://doi.org/10.1007/s00190-009-0308-3

Landerer, F. W., Flechtner, F. M., Save, H., Webb, F. H., Bandikova, T., Bertiger, W. I., et al. (2020). Extending the global mass change data record: Grace follow-on instrument and science data performance. *Geophysical Research Letters*, *47*(12), e2020GL088306. https://doi.org/10.1029/2020GL088306

Lennart, L. (1999). *System identification: Theory for the user*. PTR Prentice Hall (pp. 1–14).

Levine, P. A., Randerson, J. T., Swenson, S. C., & Lawrence, D. M. (2016). Evaluating the strength of the land−atmosphere moisture feedback in Earth system models using satellite observations. *Hydrology and Earth System Sciences*, *20*(12), 4837–4856. https://doi.org/10.5194/hess-20-4837-2016

Li, B., Rodell, M., Kumar, S., Beaudoing, H. K., Getirana, A., Zaitchik, B. F., et al. (2019). Global GRACE data assimilation for groundwater and drought monitoring: Advances and challenges. *Water Resources Research*, *55*(9), 7564–7586. https://doi.org/10.1029/2018WR024618

Lievens, H., Tomer, S. K., Al Bitar, A., De Lannoy, G. J., Drusch, M., Dumedah, G., et al. (2015). SMOS soil moisture assimilation for improved hydrologic simulation in the Murray darling basin, Australia. *Remote Sensing of Environment*, *168*, 146–162. https://doi.org/10.1016/j.rse.2015.06.025

Long, D., Scanlon, B. R., Longuevergne, L., Sun, A. Y., Fernando, D. N., & Save, H. (2013). GRACE satellite monitoring of large depletion in water storage in response to the 2011 drought in Texas. *Geophysical Research Letters*, *40*(13), 3395–3401. https://doi.org/10.1002/grl.50655

Maupin, M. A., & Barber, N. L. (2005). *Estimated withdrawals from principal aquifers in the United States, 2000*. Vol. 1279. US Department of the Interior, US Geological Survey.

Mehrnegar, N., Jones, O., Singer, M. B., Schumacher, M., Bates, P., & Forootan, E. (2020). Comparing global hydrological models and combining them with GRACE by dynamic model data averaging (DMDA). *Advances in Water Resources*, *138*, 103528. https://doi.org/10.1016/j.advwatres.2020.103528

Mehrnegar, N., Jones, O., Singer, M. B., Schumacher, M., Jagdhuber, T., Scanlon, B. R., et al. (2020). Exploring groundwater and soil water storage changes across the CONUS at 12.5 km resolution by a Bayesian integration of GRACE data into W3RA. *Science of the Total Environment*, *758*, 143579. https://doi.org/10.1016/j.scitotenv.2020.143579

Meijering, E. (2002). A chronology of interpolation: From ancient astronomy to modern signal and image processing. *Proceedings of the IEEE*, *90*(3), 319–342. https://doi.org/10.1109/5.993400

Miro, M. E., & Famiglietti, J. S. (2018). Downscaling grace remote sensing datasets to high-resolution groundwater storage change maps of California's central valley. *Remote Sensing*, *10*(1), 143. https://doi.org/10.3390/rs10010143

Muñoz Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., et al. (2019). Era5-land hourly data from 1981 to present. *Copernicus Climate Change Service (C3S) Climate Data Store (CDS)*, *10*. https://doi.org/10.24381/cds.e2161bac

O'Neill, P., Chan, S., Njoku, E., Jackson, T., Bindlish, R., Chaubell, J., & Colliander, A. (2021). Smap enhanced l3 radiometer global and polar grid daily 9 km ease-grid soil moisture, version 5, national snow and ice data center (subset of rio santa basin). *National Snow and Ice Data Center*. https://doi.org/10.5067/4DQ54OUIJ9DL

Peltier, W. R., Argus, D., & Drummond, R. (2015). Space geodesy constrains ice age terminal deglaciation: The global ice-6g_c (vm5a) model. *Journal of Geophysical Research: Solid Earth*, *120*(1), 450–487. https://doi.org/10.1002/2014JB011176

Peterson, S. M., Flynn, A. T., & Traylor, J. P. (2016). Groundwater-flow model of the northern high plains aquifer in Colorado, Kansas, Nebraska, South Dakota, and Wyoming. *Technical Report*. https://doi.org/10.3133/sir20165153

Petropoulos, G. P., Ireland, G., & Barrett, B. (2015). Surface soil moisture retrievals from remote sensing: Current status, products & future trends. *Physics and Chemistry of the Earth, Parts A/B/C*, *83*, 36–56. https://doi.org/10.1016/j.pce.2015.02.009

Rabiner, L. R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, *77*(2), 257–286. https://doi.org/10.1109/5.18626

Reichle, R. H. (2008). Data assimilation methods in the earth sciences. *Advances in Water Resources*, *31*(11), 1411–1418. https://doi.org/10.1016/j.advwatres.2008.01.001

Reichle, R. H., & Koster, R. D. (2005). Global assimilation of satellite surface soil moisture retrievals into the NASA catchment land surface model. *Geophysical Research Letters*, *32*(2), L02404. https://doi.org/10.1029/2004GL021700

Renzullo, L. J., Van Dijk, A., 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

Richard Peltier, W., Argus, D. F., & Drummond, R. (2018). Comment on "an assessment of the ice-6g_c (vm5a) glacial isostatic adjustment model" by Purcell et al. *Journal of Geophysical Research: Solid Earth*, *123*(2), 2019–2028. https://doi.org/10.1002/2016JB013844

Ridler, M.-E., Madsen, H., Stisen, S., Bircher, S., & Fensholt, R. (2014). Assimilation of SMOS-derived soil moisture in a fully integrated hydrological and soil-vegetation-atmosphere transfer model in Western Denmark. *Water Resources Research*, *50*(11), 8962–8981. https://doi.org/10.1002/2014WR015392

Sandholt, I., Rasmussen, K., & Andersen, J. (2002). A simple interpretation of the surface temperature/vegetation index space for assessment of surface moisture status. *Remote Sensing of Environment*, *79*(2–3), 213–224. https://doi.org/10.1016/S0034-4257(01)00274-7

Särkkä, S. (2013). *Bayesian filtering and smoothing*. Vol. 3. Cambridge University Press. https://doi.org/10.1201/b16018

Scanlon, B. R., Faunt, C. C., Longuevergne, L., Reedy, R. C., Alley, W. M., McGuire, V. L., & McMahon, P. B. (2012). Groundwater depletion and sustainability of irrigation in the US high plains and central valley. *Proceedings of the National Academy of Sciences*, *109*(24), 9320–9325. https://doi.org/10.1073/pnas.1200311109

Scanlon, B. R., Rateb, A., Pool, D. R., Sanford, W., Save, H., Sun, A., et al. (2021). Effects of climate and irrigation on grace-based estimates of water storage changes in major us aquifers. *Environmental Research Letters*, *16*(9), 094009. https://doi.org/10.1088/1748-9326/ac16ff

Schellekens, J., Dutra, E., Martínez-de la Torre, A., Balsamo, G., Van Dijk, A., Weiland, F. S., et al. (2017). A global water resources ensemble of hydrological models: The eartH2Observe tier-1 dataset. *Earth System Science Data*, *9*(2), 389–413. https://doi.org/10.5194/essd-9-389-2017

Schumacher, M., Forootan, E., Van Dijk, A. I. J. M., Schmied, H. M., Crosbie, R. S., Kusche, J., & Döll, P. (2018). Improving drought simulations within the murray-darling basin by combined calibration/assimilation of GRACE data into the WaterGAP global hydrology model. *Remote Sensing of Environment*, *204*, 212–228. https://doi.org/10.1016/j.rse.2017.10.029

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*(6), 537–559. https://doi.org/10.1007/s00190-016-0892-y

Smith, A. F., & Roberts, G. O. (1993). Bayesian computation via the Gibbs sampler and related Markov chain Monte Carlo methods. *Journal of the Royal Statistical Society: Series B*, *55*(1), 3–23. https://doi.org/10.1111/j.2517-6161.1993.tb01466.x

Snyder, C., Bengtsson, T., Bickel, P., & Anderson, J. (2008). Obstacles to high-dimensional particle filtering. *Monthly Weather Review*, *136*(12), 4629–4640. https://doi.org/10.1175/2008MWR2529.1

Swenson, S., Chambers, D., & Wahr, J. (2008). Estimating geocenter variations from a combination of GRACE and ocean model output. *Journal of Geophysical Research*, *113*(B8). https://doi.org/10.1029/2007JB005338

Tangdamrongsub, N., Han, S.-C., Tian, S., Müller Schmied, H., Sutanudjaja, E. H., Ran, J., & Feng, W. (2018). Evaluation of groundwater storage variations estimated from grace data assimilation and state-of-the-art land surface models in Australia and the north China plain. *Remote Sensing*, *10*(3), 483. https://doi.org/10.3390/rs10030483

Tangdamrongsub, N., Han, S.-C., Yeo, I.-Y., Dong, J., Steele-Dunne, S. C., Willgoose, G., & Walker, J. P. (2020). Multivariate data assimilation of grace, SMOS, SMAP measurements for improved regional soil moisture and groundwater storage estimates. *Advances in Water Resources*, *135*, 103477. https://doi.org/10.1016/j.advwatres.2019.103477

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*(5683), 503–505. https://doi.org/10.1126/science.1099192

Tapley, B. D., Bettadpur, S., Watkins, M., & Reigber, C. (2004). The gravity recovery and climate experiment: Mission overview and early results. *Geophysical Research Letters*, *31*(9). https://doi.org/10.1029/2004gl019920

Tapley, B. D., Watkins, M. M., Flechtner, F., Reigber, C., Bettadpur, S., Rodell, M., et al. (2019). Contributions of grace to understanding climate change. *Nature Climate Change*, *9*(5), 358–369. https://doi.org/10.1038/s41558-019-0456-2

Thelen, E., & Smith, L. B. (1998). Dynamic systems theories.

Thelin, G. P., & Heimes, F. J. (1987). *Mapping irrigated cropland from landsat data for determination of water use from the high plains aquifer in parts of Colorado, Kansas, Nebraska, New Mexico, Oklahoma, South Dakota, Texas, and Wyoming (No. 1400)*. US Government Printing Office.

Tian, S., Renzullo, L. J., Van Dijk, A. I., Tregoning, P., & Walker, J. P. (2019). Global joint assimilation of grace and SMOS for improved estimation of root-zone soil moisture and vegetation response. *Hydrology and Earth System Sciences*, *23*(2), 1067–1081. https://doi.org/10.5194/hess-23-1067-2019,2019

Tian, S., Tregoning, P., Renzullo, L. J., Van Dijk, A. I. J. M., Walker, J. P., Pauwels, V. R., & Allgeyer, S. (2017). Improved water balance component estimates through joint assimilation of GRACE water storage and SMOS soil moisture retrievals. *Water Resources Research*, *53*(3), 1820–1840. https://doi.org/10.1002/2016WR019641

Van Dijk, A. I. J. M. (2010). *The Australian water resources assessment system: Technical 901 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., 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*(8), 2955–2973. https://doi.org/10.5194/hess-18-2955-2014

Wahr, J., Molenaar, M., & Bryan, F. (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), 30205–30229. https://doi.org/10.1029/98JB02844

Xu, X., Tolson, B. A., Li, J., Staebler, R. M., Seglenieks, F., Haghnegahdar, A., & Davison, B. (2015). Assimilation of SMOS soil moisture over the Great Lakes basin. *Remote Sensing of Environment*, *169*, 163–175. https://doi.org/10.1016/j.rse.2015.08.017

Zaitchik, B. F., Rodell, M., & Reichle, R. H. (2008a). Assimilation of GRACE terrestrial water storage data into a land surface model: Results for the Mississippi River basin. *Journal of Hydrometeorology*, *9*(3), 535–548. https://doi.org/10.1175/2007JHM951.1

Zaitchik, B. F., Rodell, M., & Reichle, R. H. (2008b). Assimilation of GRACE terrestrial water storage data into a land surface model: Results for the Mississippi River basin. *Journal of Hydrometeorology*, *9*(3), 535–548. https://doi.org/10.1175/2007JHM951.1

Zribi, M., Baghdadi, N., Holah, N., & Fafin, O. (2005). New methodology for soil surface moisture estimation and its application to ENVISAT-ASAR multi-incidence data inversion. *Remote Sensing of Environment*, *96*(3–4), 485–496. https://doi.org/10.1016/j.rse.2005.04.005

Zribi, M., & Dechambre, M. (2003). A new empirical model to retrieve soil moisture and roughness from c-band radar data. *Remote Sensing of Environment*, *84*(1), 42–52. https://doi.org/10.1016/S0034-4257(02)00069-X