Reply to the Reviewers

Manuscript: egusphere-2025-2058, submitted to *HESS*

Title: **A Review of Current Best Practices and Future Directions in Assimilating GRACE/-FO Terrestrial Water Storage Data into Numerical Models**

Authors: Anne Springer, Gabriëlle De Lannoy, Matthew Rodell, Yorck Ewerdwalbesloh, Helena Gerdener, Mehdi Khaki, Bailing Li, Fupeng Li, Maike Schumacher, Natthachet Tangdamrongsub, Mohammad J. Tourian, Wanshu Nie, and Jürgen Kusche

Anonymous Referee #1: 30 Aug 2025

This paper reviews data assimilation techniques and applications that combine GRACE water storage data with hydrological modeling to improve estimation of water storages and fluxes. The paper is of interest to the HESS readership and generally informative and well written. Overall, I enjoyed reading the paper, but I think there is room for increasing impact and accessibility of the work. My comments are detailed below.

Reply: We greatly appreciate your thoughtful review and valuable feedback. Thank you for your helpful comments and suggestions, which we address point by point below. Please note that your comments are presented in italics, our replies in standard font, and suggested changes to the manuscript are highlighted in blue.

Comment 1 *Contribution of the paper: the introduction asserts that "a thorough synthesis of existing studies is not yet available" (line 119). There have been previous review papers on GRACE data assimilation, so this statement needs clarification and the specific contribution of this review paper needs to be articulated better.*

Reply: Many thanks for highlighting this issue. We agree that the addressed statement requires clarification. Previous papers have focused on specific components of the assimilation system, for example the computation and application of assimilation increments (Girotto et al., 2016) or particular applications such as improving snow estimates (Girotto et al., 2020). The dedicated review by Soltani et al. (2021) provided a method-oriented overview of how GRACE data can be assimilated into hydrological models, with a strong emphasis on error modeling and data assimilation algorithms, but also some thoughts about existing challenges. In our contribution, we aim to provide a thorough synthesis of existing studies, which have almost doubled since the 2021 review, highlighting the different areas of application and offering a systematic analysis of the applied models and the common settings used in current data assimilation frameworks. In addition, we assess the present lack of consensus within the community and outline directions that may support convergence and guide future developments. We added these aspects to the revised manuscript in line 130, in the following way:

Previous review papers have focused on providing an overview of methods for assimilating GRACE data into hydrological models, with an emphasis on error modelling and data assimilation algorithms (e.g. Soltani et al., 2021). This paper aims to provide a thorough synthesis of existing studies, highlighting the various application areas and offering a systematic analysis of the common settings within current DA frameworks. Additionally, we synthesize the current state of research, evaluate the present lack of consensus within the community regarding DA strategies, and outline directions that may support convergence and open up perspectives on new directions.

Comment 2 *Structure of the paper: the paper could benefit from a more systematic structure. For example, the authors could start by introducing the three main components of a GRACE DA system (observation model and errors, hydrological model and errors, DA algorithm and setup) and then systematically present previous work, best practices, open problems, and future directions for those* *three components. I mention these three components, because that is what the conclusions section ("synthesis" section 7) uses. The other sections of the paper do not really follow this pattern (or at least less obviously so). A more consistent structure could help readability and coherence.*

Reply: Thank you very much for your thoughts on the structure of our paper. After careful consideration, we have come to the following conclusion. We would like to keep the current structure of the manuscript that we agreed on previously. In data assimilation, the observation model is typically tailored to the filter algorithm and the hydrological model setup, or adapted/simplified as needed. A sequential description as suggested here does not fully reflect this flexibility and adaptation. Section 2 provides essential background on the models and observations forming the basis of the data assimilation approach. The GRACE observation model and errors have already been discussed in many GRACE-related and hydrology studies, e.g., Chen et al. (2022); therefore, we do not consider an in-depth discussion necessary here. As the treatment of model and observation errors is already an integral part of the assimilation strategy, we would prefer to address these aspects in Section 3, where we present the setup of the data assimilation frameworks. Furthermore, most best practices are closely linked to specific components of the data assimilation system. To avoid repetition, we therefore prefer to discuss best practices within the relevant sections, for example in the context of assimilation algorithms or observation errors. We would like to have separate sections on 'Current challenges and open issues' and 'Future directions', as we particularly want these aspects to be highlighted at a glance in our paper. Yet, we adjusted the order of "model" and "observations" in the synthesis section to ensure consistency with the structure introduced in Section 2.

Comment 3 *It's interesting that a paper on data assimilation manages to avoid any mathematical equations. On the one hand, this makes the material readable, but on the other hand it can also make things less precise/concrete. For example, section 3.5 introduces "innovations" and "increments" without defining these in an equation. The meaning of these terms is described on lines 458-460, but for readers not familiar with data assimilation the connection between them may remain a bit vague without an equation. Do the authors think this is a problem?*

Reply: We appreciate your suggestions on this topic! Indeed this comment was also made by the other reviewer. As we prefer to keep equations out of the main text, and it is not possible to include equations for all types of filters, we decided to present them in a figure. We added the key equations to the current Figure 2 of the manuscript (see Figure 1), making terms such as 'innovations' and 'increments' more accessible to the reader.

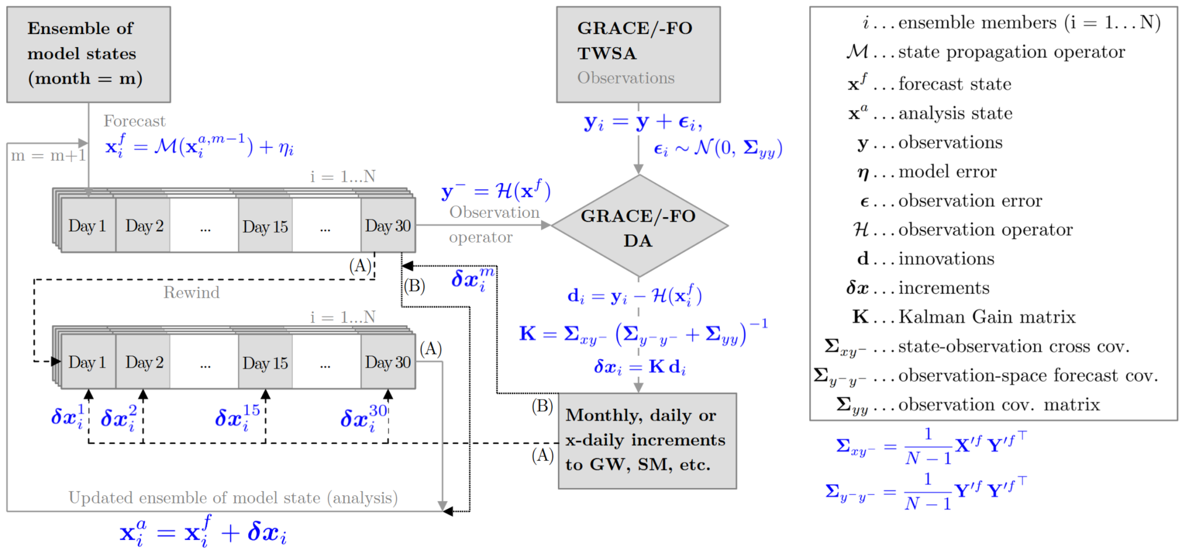

Figure 1: General concept for assimilating monthly GRACE/-FO-derived TWSA into GHMs and LSMs along with most relevant Ensemble Kalman Filter (EnKF) equations showing two options for applying the assimilation increments: After computing the increments, the model is either (A) rewound and re-run over the month with the increments $\boldsymbol{\delta x}_i^1 ... \boldsymbol{\delta x}_i^{30}$ distributed across all days, or (B) updated by applying the full monthly increment $\boldsymbol{\delta x}_i^m$ at the end of the month. Please note that the equations provided refer to the EnKF and are expressed for each ensemble member i. The Kalman Gain matrix determines the update weights based on the state and observation error covariance matrices. X'^f and Y'^f are matrices of forecast state anomalies and forecast state observation-space anomalies, respectively. Each column represents the deviation of one ensemble member from the ensemble mean. For further details on the equations for other DA algorithms, refer to the literature cited in Section 3.2.

Comment 4: Observation errors

Comment 4a *The use of grid scaling factors to restore some of the signal in GRACE data is mentioned on lines 180-181: is any advancement or alternative approach needed here since these scaling factors are derived from hydrological models and therefore contaminated by various errors (as also acknowledged on lines 224-225)?*

Reply: Thank you for this comment. We agree and now list several methods available that attempt to restore some of the lost signal in GRACE data and then explicitly refer to a method that does not depend on hydrological models. Four popular approaches that are model-dependent are:

- the multiplicative approach by Longuevergne et al. (2010)
- the additive approach by Klees et al. (2007)
- the scaling approach by Landerer and Swenson (2012)
- the unconstrained forward modeling approach by Chen et al. (2015)

To overcome dependency on hydrological models (and their associated errors),

- Vishwakarma et al. (2017) developed a data-driven approach.

We added the above references to the manuscript in l. 193 (revised manuscript l. 196) as follows:

To account for spatial leakage effects, several methods have been developed that attempt to restore the lost signal in GRACE/-FO data. Commonly used approaches that depend on hydrological model outputs are the multiplicative approach by Longuevergne et al. (2010), the additive approach by Klees et al. (2007), the grid factor scaling approach by Landerer and Swenson (2012), and the unconstrained forward modeling approach by Chen et al. (2015). Vishwakarma et al. (2017) developed a data-driven approach to overcome the dependency on hydrological models and their associated uncertainties.

Comment 4b *Line 240: why are the relationships nonlinear? It would help to specify which variables we are talking about here. For example, the relations between fine-scaled storages and fluxes (precipitation, ET...) and large-scale storage are actually linear, since both the water balance and spatial averaging are linear operations.*

Reply: We appreciate the reviewer's comment regarding the linearity of the water-balance equation and spatial averaging. In principle, if all relevant water-cycle processes were perfectly observed, the relationship between high-resolution precipitation (P), evapotranspiration (ET), and runoff (R) and large-scale water storage would be linear. However, in practice, GRACE does not observe many short-timescale fluxes or individual storage components; it only measures the net mass change integrated over large regions. Consequently, while GRACE captures the integrated mass-change signal resulting from processes such as soil-moisture dynamics, infiltration, recharge, runoff generation, groundwater–surface-water interactions, and anthropogenic water use, it does not directly observe these processes individually; their effects are embedded collectively in the coarse-scale TWS signal. These processes respond in ways that depend on the current storage state (e.g., soil moisture controls ET and infiltration rates, groundwater level controls baseflow and recharge). Because GRACE senses only the combined effect of these internal dynamics rather than each component separately, the resulting signal reflects nonlinear behavior, creating a nonlinear relationship between GRACE TWSA and fine-scale variables. Additionally, when using model-derived TWSA as a predictor, nonlinear process formulations in land-surface and groundwater modules further break linear proportionality with GRACE TWSA. Therefore, although the continuity equation itself is linear, the empirical relationship between fine-scale fluxes/model TWSA and GRACE-derived TWSA becomes nonlinear because GRACE senses only the aggregated outcome of numerous unobserved, nonlinear processes and uncertainties. This practical discrepancy motivates the use of flexible statistical and Bayesian approaches capable of representing and quantifying these nonlinear dependencies. We added this explanation in the revised manuscript in l. 267.

Comment 4c *Line 259, "DA methods take care of the horizontal and vertical downscaling by design": what if the modeling domain does not fully cover all GRACE 'grid' cells that overlap with the modeling domain?*

Reply: This can be avoided when working with spherical harmonics solutions: The GRACE/-FO data can be prepared in a way that the modelling domain and GRACE data coverage match, e.g., by computing (sub-)basin averages considering the hydrological boundaries that are also considered in the model or by computing GRACE/-FO gridded data for the specified model domain. In case of working with Mascon solutions, it is possible that the model domain and GRACE gid cells do not overlap completely, e.g., along coast lines or basin boundaries. In this case, special care has to be taken to match the model and observation domain prior to data assimilation.One approach in coastal areas is to consider only those GRACE grid cells that are covered by the model grid above a chosen percentage threshold.
We added the following clarification to l. 259 (revised document l. 295):

Prior to DA, it has to be ensured that the GRACE/-FO data coverage matches with the modeling domain.

Comment 4d *Line 281-282, "the trend of the GRACE/-FO observations is kept to correct missing trends in the model". To what extent is the trend in GRACE observations also subject to error (e.g. different GRACE products showing different trends)?*

Reply: The trend in GRACE/GRACE-FO observations can be subject to significant uncertainty, and different processing indeed can yield different trend estimates depending on region, time period, and processing choices. For example,Rodell et al. (2018) provided linear TWSA trend errors (2002-2016) derived by comparing three different Mascon solution. In many regions of the world the trend errors are between 5 and 30%, except for North America near Hudson Bay, where trend errors reached 100% probably due to the GIA correction. This global study is also confirmed by more regional studies that include different GRACE/-FO data products for the comparison of trends, e.g., Zhao and Li (2017) found 30% in the Tarim basin (2002-2015). Nonetheless, as in the data assimilation framework observation and model error is weighted against each other to derive the update state; the errors are considered for both. As future gravity missions enable the extension of the TWSA observations, it will be possible to estimates trends more certainly in the future. We added the following sentence to the manuscript to clarify this point (revised manuscript l. 318):

The GRACE/-FO trend errors are in many regions between 5–30% (or higher in regions strongly affected by GIA e.g., Rodell et al. (2018); Zhao and Li (2017)) and are then implicitly (suboptimally) considered in the DA framework by weighting both observation and model forecast error.

Comment 4e *Line 393, "uniform and uncorrelated errors": I suppose "uniform" means "spatially uniform"? Would be good to also specify which probability distribution is used for the random errors, e.g. Gaussian.*

Reply: Many thanks for pointing out this potential source of misunderstanding. We added the term spatially for clarity. You are also correct that Kalman filters typically assume Gaussian noise, which we discuss in Section 3.2 (l. 356 ff, revised manuscript: l. 400 ff).

Comment 4f *Section 3.3: this section discusses how random errors in GRACE observations are modeled, separate from systematic errors (bias) discussed in section 2.5. However, how the two are modeled (noise and bias) is linked, i.e. a modeling choice in one affects the other, so I wonder why the two are discussed separately.*

Reply: Thank you very much for your question. In this paper, we discuss bias and noise separately because bias correction is considered part of the preprocessing of observations, whereas the noise model is an integral component of the data assimilation system. To help readers understand this connection, we added a reference to Sections 3.3 and 3.4 in the Section on bias correction.

Comment 5: Forecast errors

Comment 5a *More thorough and critical review could be useful here. For example, a table that summarizes how forecast errors are computed in different GRACE DA studies (e.g. which variables are perturbed, by how much, and how; whether there is accounting for spatial/temporal correlation in forcing...). This can lead to identifying gaps and give pointers for future studies. Section 5.3 does identify some open issues with forecast errors, but it's not clear how these should be tackled. E.g. how can the suggestions in section 6.1 help with the issue of bias due to nonlinearity mentioned in section 5.3?*

Reply: Many thanks for this very good idea! We added Table 1 to summarize the perturbed state and forcing variables, along with the perturbation errors, used in several GRACE DA studies. Although the list is not exhaustive, as many studies do not provide such information, the table provides information on typical ranges of perturbation errors. Most studies account for spatial and temporal correlations, so these details are not included in the table.

Furthermore, we added a paragraph to the end of section 6.1 to discuss several approaches that can be used to address non-Gaussian related ensemble biases, as follows:

To mitigate ensemble biases associated with non-Gaussian behaviors of hydrological variables, we recommend tuning the model prior to GRACE/-FO DA so that simulated TWS is better aligned with GRACE/-FO observations. This step reduces the need for large perturbation errors, and consequently, minimizes ensemble biases. Tuning model parameters to reduce systematic errors is also critical for improving the performance of EnKF based approaches, which are designed to correct random errors rather than systematic errors; in addition, it helps reduce mass imbalances caused by large DA updates. As noted earlier, the particle filtering method (Crisan, 2001) is not restricted to specific statistical distributions and therefore can address the non-Gaussian issue; however, a large ensemble size is needed to effectively represent a highly skewed distribution. Similarly, transforming functions have been used for non-Gaussian data assimilation within a 3-D variational method (Van Loon and Fletcher, 2023) and their application in variants of the EnKF warrants future investigation.

Table 1: Summary of perturbations applied in various hydrological and land surface models for GRACE/-FO DA, including the types of state variables, meteorological forcing variables, and calibrated parameters perturbed. CD: Catchment deficit, SE: Surface excess, SMC: soil moisture; SWE: snow water equivalent; GWS: groundwater storage; PCP: precipitation; SWR: shortwave radiation; LWR: longwave radiation; RAD: radiation; TEMP: air temperature; minT/maxT: minimum/maximum air temperature; PET: potential evapotranspiration; infR: infrared surface temperature; SysE: systematic error; radE: random error; TRMM: Tropical Rainfall Measuring Mission; maxDrain: maximum subsurface drainage rate; DrainRt: drainage rate; TWS: terrestrial water storage, see Table 1 in the main manuscript for model abbreviations.

Model	**Perturbed state variable**	**Perturbed meteorological forcing variable**	**Perturbed parameters (input or calibrated)**	**Reference**
AWRA-L	–	PCP (50%), RAD (30%), minT (0.3°), maxT (0.25°)	–	Shokri et al. (2018)
CABLE	–	SWR (10%), TEMP (10%), PCP (TRMM product)	Soil texture (10%), Saturated fraction (10%), maxDrain (10%), DrainRt (10%)	Tang et al. (2020)
CLM	–	PCP (30%), SWR (30%), LWR (30 W/m2), TEMP (2°)	Soil texture (10%)	Su et al. (2010); Springer et al. (2019)
CLM-ParFlow	–	PCP (10%)	Soil texture (10%)	Soltani et al. (2024)
CLSM	CD (0.02 mm), SE (0.05 mm), SWE (0.12%)	PCP (50%), SWR (30%), LWR (50 W/m2)	–	Reichle et al. (2007); Girotto et al. (2016)
MESH	SWE (0.0004 mm)	PCP (50%), SWR (30%), LWR (20 W/m2)	–	Bahrami et al. (2021)
MGB	–	PCP (25% SysE and 70% radE)	Several storage, residence time and river related parameters	Wongchuig et al. (2024)
Noah	–	PCP, TEMP, RAD, infR, TEMP (1% of mean)	–	Liu et al. (2021)
Noah-MP	SMC (10^{-5}-10^{-4}mm3/mm3), GWS (0.01 mm)	PCP (30%), SWR (30%), LWR (50 W/m2)	–	Nie et al. (2019)
PCR-GLOBWB	–	TEMP (2°), PET (30%), PCP (TRMM product)	15 TWS-related parameters (20%)	Tangdamrongsub et al. (2017)
wflow_hbv	–	PCP (10%), TEMP (15%), PET (15%)	SMC and runoff routine parameters (10%)	Tangdamrongsub et al. (2015)
WGHM	–	PCP (30%), TEMP (2°)	22 parameters	Eicker et al. (2014); Schumacher et al. (2016)
W3RA	–	PCP (60%), SWR (50 W/m2), TEMP (2°)	–	Tian et al. (2017)

Comment 5b *It seems generating the forecast ensemble is mostly done offline, e.g. the authors mention sensitivity analysis. Are there opportunities for calibrating forecast uncertainties as part of the DA system, i.e. in an automated fashion?*

Reply: Thank you for the comment. Indeed, modern DA systems also provide opportunities to calibrate forecast uncertainties automatically during runtime. We added some more clarification to Section 3.4 to address this, as following:

Furthermore, modern DA systems also provide opportunities to calibrate forecast uncertainties automatically during runtime, rather than relying solely on offline perturbation design. Techniques such as adaptive covariance inflation (Anderson, 2007) and relaxation-to-prior-spread (RTPS) or relaxation-to-prior-perturbations (RTPP) (Whitaker and Hamill, 2012) dynamically adjust ensemble spread using innovation statistics to compensate for under- or over-dispersive ensembles. In addition, stochastic model error estimation and hierarchical Bayesian approaches allow perturbation magnitudes or model error parameters to be updated online (Ruiz et al., 2013; Berry and Harlim, 2017). These automated strategies reduce the reliance on extensive offline sensitivity analyses and enable forecast uncertainties to evolve consistently with model–observation discrepancies.

Comment 6: DA algorithm and setup

Comment 6a *The paper would benefit from more clearly discussing the relation between the intended goal of the DA application and how DA is implemented. The first sentence of the abstract suggests that the main goal here is reanalysis, i.e. create consistent historical datasets of water storage and fluxes that incorporate information from GRACE. However, it seems most of the paper discusses studies and implementations that are more related to operational DA (for use in e.g. early drought warning systems), as evidenced by the focus on filtering implementations (or smoothing implementations that only look at the last month) where DA is used to update initial storages for the next forecast. For reanalysis purposes it seems more appropriate to use smoothing implementations that make use of the entire historical record. Do the authors agree with this? If so, it would be helpful to include papers that use smoothing (not just last month) to assimilate GRACE data.*

Reply: We agree with the reviewer. For reanalysis purposes, it would be more appropriate to use a smoothing implementation that crosses multiple months, if we wanted to focus on long-term variability. However, to our knowledge, the few existing land re-analysis products have traditionally always used a 'filter-like' approach. This is because TWS varies at shorter timescales and because of technical limitations: smoothing over the entire historical record would require a lot of memory, and doing so for continental or global applications would be prohibitive.
We updated the text as follows:

L.9: With processing pipelines now being developed for low-latency short-term data products from the upcoming next-generation gravity missions, we expect that low-latency periodically updated reanalyses and analyses from assimilation will become more relevant.
L.43 (revised document l. 49): DA can be used to produce long-term reanalysis estimates of TWS to support land system understanding (Baatz et al., 2021), or obtain the best current state estimate for operational forecasting, as is needed for early warning systems (e.g. drought). Given the monthly resolution of GRACE/-FO observations, the line between both is vague, and we loosely use the term reanalysis for both.
L.296 (revised document l. 334): All currently available continental to global-scale land reanalysis products typically use this 'filter-like' approach, i.e. using a filter or a smoother with non-overlapping short one-month windows, for computational efficiency and because TWS varies at timescales of less than a month. For long-term reanalysis, longer or moving smoothing windows could be explored in the future, as is done for atmospheric or oceanographic reanalyses.

Comment 6b *Likewise, when the aim is in generating consistent estimates of the various water storages and fluxes, it is interesting that the paper reports (line 669) that most DA studies violate the water balance (which clearly introduces inconsistency in the estimates). I guess this relates to which variables are being updated by the DA system. Most DA studies cited in the paper only update the storages, but there are various studies that use Kalman filtering and smoothing techniques to update other variables of the water balance as well with the aim of maintaining a closing water balance.*

Reply: Exactly, we point out in line 669 (new manuscript l. 736) that most GRACE DA implementations violate the water balance. As the reviewer notices, this is related to the fact that most GRACE DA studies update the water storages, but also since the vast majority of GRACE DA implementations rely on ensemble filters (as opposed to variational approaches). It is a well-known property of ensemble Kalman filter DA that mass conservation is violated, not only in case of GRACE DA but also for other assimilated variables. We would be glad to cite papers here that overcome this problem but in fact we do not know of any, and we appreciate if the reviewer could provide their suggestions. On the other hand, the suite of data sets used in the model/data system may very well violate the water balance from the start, due to unavoidable data errors, and some degree of violation cannot be avoided at all while making use of the data. To clarify this, we expanded our statement in the manuscript to

...violates the water balance (...) to some extent, which is the price to be paid to nudging a model run closer to real observations. Techniques exist to mitigate this effect.

Comment 6c *Line 357, "the EnKF and EnKS are optimal and unbiased only when assuming Gaussian errors": these algorithms are 'optimal' when you assume the errors are Gaussian or when that assumption is correct?*

Reply: Thanks for asking for clarification! These algorithms are statistically 'optimal' when the assumption is correct, i.e. the errors are Gaussian. We modified the sentence to

Furthermore, strictly speaking the EnKF and EnKS are optimal and unbiased only when errors are Gaussian; however unless strong nonlinearities lead to violations of this assumption they usually work in a satisfactory way.

Comment 7 *Abstract: would benefit from a rewrite, as it currently reads more like an introduction and does not contain concrete information on the findings ("best practices and future directions" as promised by the title of the paper).*

Reply: We get your point! We rewrote the second part of the abstract addressing best practices and future direction as following:

The review reveals that the most effective assimilation strategies leverage (robust modifications of) the classical ensemble Kalman filter and localization techniques, explicitly account for correlated observation errors, and address biases contained in the observations as well as those arising from model perturbations. Unmodeled processes must be carefully handled through signal separation, multisource assimilation, or removal prior to assimilation. Future directions include developing low-latency products for near-real-time assimilation, integrating enhanced and combined satellite observations, and employing machine-learning approaches for downscaling and hybrid assimilation. Collectively, these strategies provide a pathway toward more accurate, physically consistent, and operationally useful water cycle reanalyses.

Comment 8 *Conclusions: several open issues identified in the text don't find their way into the conclusions.*

Reply: We are not entirely sure which specific point the reviewer is addressing. We assume it relates to Section 7, the Synthesis, but from the comment it is not entirely clear what the reviewer considers to be missing. In this section, our goal was to summarize the most important open issues that we believe should be highlighted and addressed next. In particular, we explicitly discuss:

- Physical realism and representation of anthropogenic processes in models,
- Minimizing water budget imbalances caused by DA,
- Model forecast uncertainty,
- Bias correction and correction of geophysical signals,
- Observation model error and future gravity products,
- Spatial resolution mismatch, localization, and handling of spatial correlations,
- State vector setup and computation of innovations,
- Application of assimilation increments, and
- Handling of temporally downscaled products, as well as strategies for systematic validation of DA experiments.

We deliberately do not discuss in detail issues such as ensemble generation methods, as we aimed to keep the synthesis focused on the most pressing and broadly relevant topics. However, we added one important aspect to the first bullet point of the synthesis:

Most current GRACE–FO DA systems rely on Gaussian assumptions, which can restrict the representation of skewed or heavy tailed uncertainties in water storage dynamics, so future work should develop more flexible assimilation approaches that allow advanced statistical descriptions of errors, including non Gaussian methods.

Comment 9 *Edits: (line 30) at the other hand –¿ on the other hand, (line 381) weighing –¿ weighting, (line 247) GARCE –¿ GRACE.*

Reply: Applied - thanks for your attention!

Anonymous Referee #2: 14 Oct 2025

General comment: In this manuscript, the authors provide a comprehensive review of current practices and future directions in assimilating terrestrial water storage (TWS) data from the GRACE and GRACE-FO missions into hydrological and land surface models. It draws on approximately 200 references to highlight advancements in frameworks like the Ensemble Kalman Filter (EnKF) and its variants, while addressing practical issues such as scale mismatches, error correlations, and applications in drought monitoring and climate trend analysis. The emphasis on geophysical corrections (e.g., glacial isostatic adjustment) and the integration of multi-sensor data (e.g., with SMAP or SWOT) provides a cohesive narrative that bridges theoretical DA with operational hydrology. The manuscript excels in its comprehensive scope and accessibility, making it a valuable resource for early-career researchers and practitioners seeking an entry point into GRACE/-FO DA. It includes detailed tables (e.g., Table 1: models that evaluated TWS; Table 2: DA frameworks) and illustrative figures (e.g., Figure 1: TWS components scheme; Figure 3: statistics from evaluations made by models).
The manuscript is timely, given the impending launch of missions like GRACE-C (2030) or NGGM and advancements in Earth System Modeling (ESM) frameworks. In general, this work is worth publishing in the Hydrology and Earth System Sciences journal, as it fills a gap in synthesizing post-GRACE-FO literature and could serve as a foundational reference for advancing standardized practices in hydrological DA, however some improvements are necessary, including deeper critical analysis and enhanced didactic elements. Some suggestions for revisions are included below and these points do not detract from its value; rather, they could enhance its rigor and relevance.

Reply: We sincerely thank you for your kind and appreciative words, as well as for the time and effort you have devoted to reviewing our work. We address your suggestions in detail below. Please note that your comments are presented in italics, our replies in standard font, and suggested changes to the manuscript are highlighted in blue.

Moderate comments: *While the manuscript is a solid synthesis, it requires moderate revisions to elevate its critical depth and utility. As a review paper, it should compile and critically evaluate the literature, identifying inconsistencies and unresolved debates. At present, the discussion is somewhat descriptive, with limited critique of methodological limitations or comparative assessments. To ensure the article meets the standards for publication, the authors must address the following critical points:*

Reply: Thank you for pointing this out. Your understanding precisely reflects what we intended with this review. We are grateful for your thoughtful ideas and hope that our revisions further clarify our intentions.

Comment 1 *Enhance your critical evaluation and gap analysis. In order to avoid a mere listing of works, I suggest to add a dedicated sub-section in Section 5 to quantify gaps. One option would be to use meta-analysis of DA performance metrics from recent studies (2020-2025).*

Reply: The meta-analysis is an interesting suggestion. However, we believe applying meta analysis to studies that use all kinds of different assumptions, error models, hydrology models and perturbations, and finally are geared towards differing applications and metrics, is a fallacy. As we suggest in this paper we need to introduce a benchmarking approach first before we can apply meta analysis to the results of different studies. We hope that this addresses the comment, which was not very clear to us. Regarding a more critical evaluation and comparison of different DA approaches, we included a Table comparing strengths and weaknesses of different assimilation approaches (Table 2).

Comment 2 *Improve validation and uncertainty discussions in Section 4, since it lacks depth in metrics; mandate inclusion of quantitative benchmarks (e.g., Triple Collocation Analysis) and error propagation models. I suggest, for instance, to address how non-Gaussianity affects DA reliability, proposing solutions to make recommendations more actionable.*

Reply: We appreciate the reviewer's comment. We expanded the discussion in Section 4 to provide a more comprehensive treatment of validation metrics and uncertainty aspects. In particular, we added an explanation of Triple Collocation, emphasizing its usefulness when multiple independent datasets are used to validate a single variable. We also extended the section on uncertainty treatment by including a discussion on error propagation and now briefly mention error propagation models. Furthermore, we added a statement on non-Gaussianity, explaining that non-Gaussian error distributions can degrade filter performance and validation metrics that assume Gaussianity. We now recommend employing information-theoretic measures in such cases to ensure robust validation.

We updated the text as follows:

L.562 (revised manuscript l. 615): Remotely sensed hydrological products are now widely used to validate GRACE/-FO DA simulations (Kumar et al., 2016; Zhao and Yang, 2018; Khaki et al., 2017; Jung et al., 2019). However, these datasets often differ in accuracy and consistency, making robust comparisons challenging (van Dijk et al., 2014). When multiple independent datasets are available for the same hydrological variable, triple collocation provides a way to assess the quality of each dataset without assuming that any of them is error-free (Stoffelen, 1998; Gruber et al., 2016). The method separates unpredictable measurement noise (random errors) from systematic offsets or scaling differences (calibration biases) based solely on the mutual agreement among the datasets. As such, triple collocation offers a practical solution for validation when more than two datasets are collocated in space and time and should therefore be considered a standard option.
L.573 (revised manuscript l. 629): Moreover, any preprocessing operations applied to the data (for example deseasonalizing, detrending, spatial averaging or temporal aggregation) introduce additional uncertainty which must be propagated into the final validation metrics. Error propagation models can

be used to quantify the uncertainty of observational datasets by estimating how input measurement errors or retrieval assumptions propagate into derived variables. Such approaches can provide spatially explicit uncertainty estimates for each observation, thereby offering a more comprehensive view of dataset reliability than local in-situ validations alone (Dorigo et al., 2010). Neglecting these aspects can lead to overconfidence in skill estimates or misinterpretation of variance metrics, particularly when multiple data sources or preprocessing steps are involved.
An additional challenge is that both the model state distributions and even the retrieval error distributions may depart significantly from Gaussian assumptions. If validation metrics assume Gaussian error or linear relationships (e.g. RMSE, linear regression), the non-Gaussian nature of the underlying distributions may skew results or invalidate confidence statements. We therefore recommend assessing the distributional characteristics of the errors (e.g., skewness) and, where needed, applying non-Gaussian or rank-based metrics (e.g., information-theoretic scores (Kumar et al., 2018; Maina et al., 2024)) to ensure robust validation.

Comment 3 *Add didactic elements to enhance educational value, incorporate summary equations for core DA techniques in Section 3. I suggest adding schematic diagrams, such as a flowchart comparing sequential vs. smoothing DA methods, or a matrix illustrating error sources in GRACE processing.*

Reply: Thank you very much for these suggestions. This comment is also related to the remark in subsection 2.5. We followed your recommendation and incorporated the summary equations into the flow chart of the DA concept (Figure 2 in the manuscript, see Figure 1). However, we prefer not to include an illustration of the GRACE error sources, as these are not the primary focus of our paper and have been discussed by many previous studies. Instead we added a reference to a paper describing details of the GRACE processing and also error sources, as follows in Line 171 (revised manuscript l. 184):

These monthly gravity field estimates are typically accompanied by an error estimate of the measurement noise, provided either as formal errors or as full covariance matrices (Chen et al., 2022).

Specific comments (Line-by-line comments):

Abstract: *1-16: The abstract is concise but could better emphasize the synthesis's novelty (e.g., "This review synthesizes over "n" studies to propose unified best practices..."). Add a sentence on key gaps identified, such as non-Gaussian challenges.*

Reply: Many thanks for your suggestions. We revised the second part of the abstract to more clearly highlight best practices and future directions, which will also implicitly address some of the identified gaps. Additionally, we incorporated your suggestion regarding the novelty of this review. We suggest the following modification to the second part of the abstract:

In light of the upcoming launches of next-generation gravity missions and the development of increasingly sophisticated Earth system modeling frameworks, this review synthesizes insights from approximately 60 GRACE/-FO data assimilation studies in an attempt to converge to best practices. The review reveals that the most effective assimilation strategies leverage (robust modifications of) the classical ensemble Kalman filter and localization techniques, explicitly account for correlated observation errors, and address biases contained in the observations as well as those arising from model perturbations. Unmodeled processes must be carefully handled through signal separation, multisource assimilation, or removal prior to assimilation. Future directions include developing low-latency products for near-real-time assimilation, integrating enhanced and combined satellite observations, and employing machine-learning approaches for downscaling and hybrid assimilation. Collectively, these strategies provide a pathway toward more accurate, physically consistent, and operationally useful water cycle reanalyses.

Introduction: *Clarify the distinction between GRACE and GRACE-FO data continuity; reference Vishwakarma et al. (2021) for gap-filling techniques. I suggest to add a simple timeline figure of GRACE missions to make the historical context more didactic.*

Reply: We thank the reviewer for highlighting the importance of clarifying GRACE and GRACE-FO continuity. In the revised manuscript, we now explicitly discuss the observational gap and reference existing gap-filling methods, including hydrological models, autoregressive approaches, and machine learning reconstructions. We prefer to do this in Section 2 rather than in the introduction, so that the introduction remains concise and focused on only the most relevant information. We also note that we could not identify any peer-reviewed publication by Vishwakarma et al. (2021) that addresses the GRACE–GRACE-FO gap; their work appears restricted to downscaling applications. We modified the beginning of Section 2.2 to:

The GRACE satellite mission (Tapley et al., 2004) monitored global TWS changes (Wahr et al., 1998) from 2002 to 2017, with its follow-on mission GRACE-FO, launched in 2018 (Landerer et al., 2020), continuing these observations. TWS changes are an essential climate variable reflecting the impact of global climate change on water resources (Rodell and Reager, 2023).

Furthermore, we added details regarding gap filling techniques:

Although GRACE and GRACE-FO provide a largely continuous record of TWSA, there is an observational gap between the missions of almost one year. Several studies have addressed this gap using different methods, including physically based hydrological models (Zhang et al., 2022), interpolation and statistical approaches such as singular spectrum analysis (SSA) and autoregressive (AR) models (Lecomte et al., 2024; Lenczuk et al., 2022), and machine learning techniques including convolutional or recurrent neural networks (Uz et al., 2022; Mo et al., 2022). Dynamic mode decomposition (Libero and Ciriello, 2025) of GRACE/-FO data to extract essential spatial and temporal patterns could also support these efforts. These techniques allow reconstruction of TWSA during the gap while preserving both seasonal cycles and long-term trends.

Sections 2 to 7:

Sub-section 2.3 *Expand on geophysical corrections with an equation for leakage error correction.*

Reply: In l. 223-225 (original manuscript), we write: "To account for leakage globally before DA, rescaling is commonly applied, which means that TWSA scale factors are estimated from hydrological model output, which in turn creates another source of uncertainty."
In fact, there are several approached how to restore the GRACE/-FO signal due to signal damping in the filtering process. We reformulated the sentence to (see revised manuscript l. 248):

To account for leakage before DA, approaches based on hydrological model output or data-driven estimates are available, but these could in turn also create another source of uncertainty (see Sect. 2.2 for details).

Furthermore, we added the named approaches in l. 196 of the revised manuscript to provide a fuller picture of the available methods:

To account for spatial leakage effects, several methods have been developed that attempt to restore the lost signal in GRACE/-FO data. Commonly used approaches that depend on hydrological model outputs are the multiplicative approach by Longuevergne et al. (2010), the additive approach by Klees et al. (2007), the grid factor scaling approach by Landerer and Swenson (2012), and the unconstrained forward modeling approach by Chen et al. (2015). Vishwakarma et al. (2017) developed a data-driven approach to overcome the dependency on hydrological models and their associated uncertainties.

It is beyond the scope of this paper, to provide equations for all of the named approaches and the reader is referred to the cited literature for this purpose.

Sub-section 2.4 *Include a recent reference (Gerdener et al., 2024) on GRACE/-FO data uncertainties. Propose a schematic diagram illustrating the GRACE processing chain from raw measurements to TWS anomalies.*

Reply: Thanks, we included the reference for the GRACE/-FO data uncertainties in this section and added a sentence. The processing chain is indeed an interesting part. However, most systems include level-2 measurement and only few use level-1 measurement but none uses raw measurements. In addition, the processing, e.g., from level 2 spherical harmonic to level-3 TWS anomalies changes from assimilation system to assimilation system. Due to these reasons and due to the fact that figures of processing schemes already widely exist, we would like to refer to existing illustrations here and keep the description of the levels that we provided in the section of GRACE/-FO data products as is. These illustrations contain detailed information of, for example, processing from level 0 raw measurements to level 2 spherical harmonics[1][2] (e.g., Ferreira et al., 2023; Gerdener, 2024).

Sub-section 2.5 *Although the article is a review paper on the application of data assimilation techniques, mainly from the Kalman filter family, I recommend at least adding the basic EnKF equations for accessibility.*

Reply: We appreciate your suggestions on this topic! Indeed this comment was also made by the other reviewer. As we prefer to keep equations out of the main text, and it is not possible to include equations for all types of filters, we decided to present them in a figure. We added the key equations to the current Figure 2 of the manuscript (see Figure 1), making terms such as 'innovations' and 'increments' more accessible to the reader.

Sub-section 3.2 *I strongly suggest a table or matrix comparing DA methods' computational costs and assumptions.*

Reply: Many thanks for this suggestion, which we are happy to follow. We added a table to this section to clarify the computational costs, relative strengths, and limitations of the DA methods discussed (Table 2):

[1] `https://www2.csr.utexas.edu/grace/asdp.html`, last accessed 11.11.25

[2] `https://www.gfz.de/en/section/global-geomonitoring-and-gravity-field/projects/closed-projects/grace-gravity-recovery-and-climate-experiment-mission/grace-products`, describe how TWS anomaly grid is derived from the level 2 spherical harmonics (e.g., Feng, 2019) or provide an overview of the common levels (within levels 0 to 3) plus integration of other approaches for example, assimilation[3], last accessed 11.11.25

Table 2: Comparison of DA methods and ensemble tuning techniques commonly applied in GRACE/GRACE-FO DA cases. Computational cost is rated qualitatively: ★☆☆☆☆ (low), ★★☆☆☆ (moderate), ★★★☆☆ (moderate–high), ★★★★☆ (high), ★★★★★ (very high).

Method / Technique	**Computational Cost**	**Key Assumptions**	**Strengths**	**Limitations / Notes**
EnKF	★★☆☆☆ (moderate)	Assumes Gaussian unbiased errors; ensemble approximates covariance	Efficient for large-scale hydrology; real-time capable; uncertainty quantification	Sampling errors for small ensembles; prone to ensemble collapse; requires localization and inflation
EnKS	★★★☆☆ (moderate–high)	Same as EnKF; assumes temporal error correlation over smoothing window	Improved temporal consistency; suitable for monthly GRACE/-FO assimilation windows	Higher memory and computational cost; retrospective only
Square-root filters (ETKF, EAKF, ESTKF)	★★☆☆☆ (moderate)	Same Gaussian assumptions as EnKF	Uncertainty quantification, reduced sampling noise; improved numerical stability; smaller ensembles possible	Still requires localization and inflation; additional linear algebra steps
Particle Filters / Particle Smoothers	★★★★★ (very high)	No Gaussian assumption; fully general Bayesian formulation	Handles nonlinear and non-Gaussian dynamics; accommodates multimodal states	Computationally infeasible for global LSMs; particle degeneracy unless many particles are used
Hybrid Machine Learning–DA	★☆☆☆☆ – ★★★☆☆ (variable)	Requires representative training data; assumes model transferability	Can reduce structural model error; enables downscaling or emulation; useful where physics are weak	Can generalize poorly; GRACE/-FO record is short; physical consistency may weaken
Localization (covariance or domain)	★★☆☆☆ – ★★★☆☆	Forecast error correlations decay with distance	Mitigates spurious long-range correlations; essential for small ensembles; improves stability in GRACE/-FO DA	Choice of radius is subjective; too strong localization distorts updates; often combined with inflation
Covariance Inflation	★☆☆☆☆ (low)	Forecast ensemble under-dispersion can be corrected multiplicatively or additively	Prevents filter divergence; widely used in GRACE/-FO DA; prevents underestimation of forecast model error covariance	Over-inflation causes noise or instability; parameters often tuned empirically
Adaptive Inflation	★★☆☆☆ (moderate)	Innovation statistics correspond to correct ensemble spread	Online tuning reduces need for offline sensitivity tests; improves filter robustness	Additional computation; may react poorly to biased observations or large model error

Sub-section 6.1 *Strengthen ML discussion with a hybrid DA-ML schematic (e.g., EnKF with LSTM for error modeling).*

Reply: Indeed, this is an important point, which was missing in the manuscript. Many thanks for this suggestion! We added to Section 6.1 the following text and Figure:

Recent years have seen substantial progress in applying ML within DA systems for numerical models, particularly in geophysical applications such as improving regional climate simulations (Bocquet, 2023; He et al., 2023; Keller and Potthast, 2024). GRACE/GRACE-FO DA frameworks can benefit from ML at several stages of the assimilation workflow (Figure 2). As described above, ML can support observation preprocessing, including gap filling and spatial downscaling of coarse TWSA fields. ML can potentially act as a surrogate observation operator and might also support the model update step by learning hydrologically meaningful distributions of DA increments across storage compartments. More intrusive uses involve learning from the state corrections produced during assimilation, enabling targeted model-error correction at each forecast step (Arcucci et al., 2021). Finally, the DA algorithms themselves (e.g., EnKF variants) can benefit from ML – for instance, by automatically correcting the ensemble spread through ML-based adaptive inflation or improving covariance structure by ML-based localization schemes – and may even be replaced by ML architectures designed to emulate full sequential filtering. However, these potential benefits need to be balanced against the substantial computational cost of training such models. In several meteorological applications, the training effort can exceed the runtime savings unless transfer learning strategies are used to adapt pre-trained models to new domains or conditions. Similar considerations are expected for GRACE/GRACE-FO DA, where the limited length of the observational record poses additional challenges for generating suitable training data.

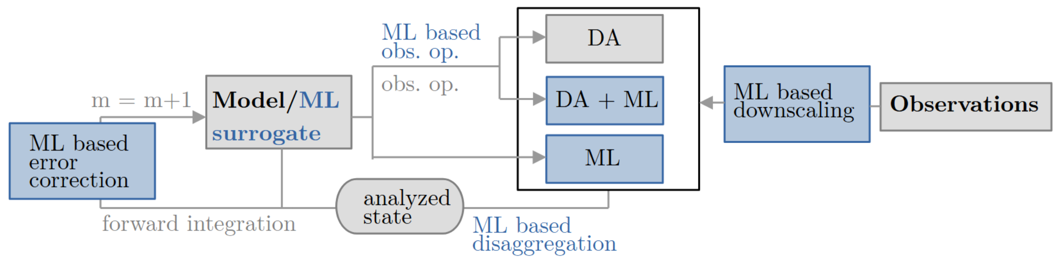

Figure 2: Workflow of a data assimilation framework with potential machine-learning augmentations highlighted in blue; note that these enhancements represent alternative components of the workflow, and it is generally not meaningful to apply ML at all components simultaneously.

References and General: *Ensure all citations are up-to-date, specially from 2020-2025. The manuscript could benefit from an appendix with a glossary of DA terms for non-experts.*

Reply: This is a very good idea! We added a glossary to the appendix (see Table below). Furthermore, we added some references regarding gap filling approaches and the following recent studies that focus on GRACE/-FO DA:
Retegui-Schiettekatte et al. (2025)
Arciniega-Esparza et al. (2025)

Table 3: Glossary of key data assimilation terms used in this study.

Term / Acronym	Definition
3D-Var	Three-dimensional variational data assimilation: adjusts the model state at a single time step by minimizing a cost function combining background and observations.
4D-Var	Four-dimensional variational data assimilation: adjusts initial conditions over a time window to match observations distributed in time, minimizing a cost function subject to model dynamics.
Analysis state	The model state after assimilation, obtained by optimally combining the background state and observations.
Background state	The model state prior to assimilation, representing the best estimate of the system before incorporating observations.
Bias correction	Adjustments applied to model or observations to remove systematic differences.
DA	Data Assimilation: combining model predictions and observations while accounting for uncertainties.

Term / Acronym	**Definition**
EnKF	Ensemble Kalman Filter: sequential DA using an ensemble to estimate state and error covariances.
EnKS	Ensemble Kalman Smoother: updates states over a time window using past and future observations to improve earlier estimates. For GRACE/GRACE-FO TWSA assimilation, the smoothing window is typically one month, meaning the monthly observation is used to redistribute the increment across the days of that month for improved temporal consistency.
Error covariance matrix	Represents uncertainties in the background model and/or observations; used to weight contributions in DA.
Filter	Sequential DA method updating the state at observation times.
Hybrid methods	Data assimilation approaches combining ensemble-based and variational techniques to leverage the advantages of both.
Increment	Adjustment applied to the model state computed by the DA algorithm.
Innovation	Difference between observations and model-predicted observations.
Inflation	A technique to artificially increase the ensemble spread in sequential DA, compensating for underestimation of uncertainties due to finite ensemble size or model errors.
Localization	Limits the influence of observations to nearby model grid points to reduce spurious correlations.
Observation operator (H)	Maps model state variables to observation space for comparison with observations.
Particle Filter (PF)	Nonlinear, non-Gaussian DA method using a weighted ensemble (particles) to represent the probability distribution of the state.
Rewind / Re-run	Strategy distributing monthly increments across previous days to improve temporal consistency.

References

Anderson, J. L.: An adaptive covariance inflation error correction algorithm for ensemble filters, Tellus A: Dynamic Meteorology and Oceanography, 59, 210–224, https://doi.org/10.1111/j.1600-0870.2006.00216.x, 2007.

Arciniega-Esparza, S., Hernández-Espriú, J. A., Salinas-Calleros, G., Birkel, C., and Sanchez, R.: Assessing hydrological drought propagation through assimilation of GRACE for groundwater storage anomalies modelling in northeastern Mexico, Journal of Hydrology, 661, 133 826, https://doi.org/10.1016/j.jhydrol.2025.133826, 2025.

Arcucci, R., Zhu, J., Hu, S., and Guo, Y.-K.: Deep Data Assimilation: Integrating Deep Learning with Data Assimilation, Applied Sciences, 11, 1114, https://doi.org/10.3390/app11031114, 2021.

Baatz, R., Hendricks Franssen, H. J., Euskirchen, E., Sihi, D., Dietze, M., Ciavatta, S., Fennel, K., Beck, H., De Lannoy, G., Pauwels, V. R. N., Raiho, A., Montzka, C., Williams, M., Mishra, U., Poppe, C., Zacharias, S., Lausch, A., Samaniego, L., Van Looy, K., Bogena, H., Adamescu, M., Mirtl, M., Fox, A., Goergen, K., Naz, B. S., Zeng, Y., and Vereecken, H.: Reanalysis in Earth System Science: Toward Terrestrial Ecosystem Reanalysis, Reviews of Geophysics, 59, e2020RG000 715, https://doi.org/10.1029/2020RG000715, 2021.

Bahrami, A., Goïta, K., Magagi, R., Davison, B., Razavi, S., Elshamy, M., and Princz, D.: Data assimilation of satellite-based terrestrial water storage changes into a hydrology land-surface model, Journal of Hydrology, 597, 125 744, https://doi.org/10.1016/j.jhydrol.2020.125744, 2021.

Berry, T. and Harlim, J.: Correcting biased observation model error in data assimilation, Monthly Weather Review, 145, 2833–2853, https://doi.org/10.1175/MWR-D-16-0428.1, 2017.

Bocquet, M.: Surrogate modeling for the climate sciences dynamics with machine learning and data assimilation, Frontiers in Applied Mathematics and Statistics, 9, https://doi.org/10.3389/fams.2023.1133226, 2023.

Chen, J., Wilson, C., Li, J., and Zhang, Z.: Reducing leakage error in GRACE-observed long-term ice mass change: a case study in West Antarctica, J Geod, 89, 925–940, https://doi.org/10.1007/s00190-015-0824-2, 2015.

Chen, J., Cazenave, A., Dahle, C., Llovel, W., Panet, I., Pfeffer, J., and Moreira, L.: Applications and Challenges of GRACE and GRACE Follow-On Satellite Gravimetry, Surveys in Geophysics, 43, 305–345, https://doi.org/10.1007/s10712-021-09685-x, 2022.

Crisan, D.: Particle Filters — A Theoretical Perspective, in: Sequential Monte Carlo Methods in Practice, edited by Doucet, A., Freitas, N., and Gordon, N., pp. 17–41, Springer New York, New York, NY, ISBN 9781441928870 9781475734379, https://doi.org/10.1007/978-1-4757-3437-9_2, 2001.

Dorigo, W. A., Scipal, K., Parinussa, R. M., Liu, Y. Y., Wagner, W., de Jeu, R. a. M., and Naeimi, V.: Error characterisation of global active and passive microwave soil moisture datasets, Hydrology and Earth System Sciences, 14, 2605–2616, https://doi.org/10.5194/hess-14-2605-2010, publisher: Copernicus GmbH, 2010.

Eicker, A., Schumacher, M., Kusche, J., Döll, P., and Schmied, H. M.: Calibration/data assimilation approach for integrating GRACE data into the WaterGAP Global Hydrology Model (WGHM) using an ensemble Kalman filter: First results, Surveys in Geophysics, 35, 1285–1309, 2014.

Feng, W.: GRAMAT: A comprehensive Matlab toolbox for estimating global mass variations from GRACE satellite data, Earth Science Informatics, 12, 389–404, 2019.

Ferreira, V., Yong, B., Montecino, H., Ndehedehe, C. E., Seitz, K., Kutterer, H., and Yang, K.: Estimating GRACE terrestrial water storage anomaly using an improved point mass solution, Scientific Data, 10, 234, 2023.

Gerdener, H.: A global drought monitoring framework using GRACE/-FO data assimilation, https://doi.org/10.48565/BONNDOC-438, 2024.

Girotto, M., De Lannoy, G. J., Reichle, R. H., and Rodell, M.: Assimilation of gridded terrestrial water storage observations from GRACE into a land surface model, Water Resources Research, 52, 4164–4183, 2016.

Girotto, M., Musselman, K. N., and Essery, R. L. H.: Data Assimilation Improves Estimates of Climate-Sensitive Seasonal Snow, Current Climate Change Reports, 6, 81–94, https://doi.org/10.1007/s40641-020-00159-7, 2020.

Gruber, A., Su, C. H., Zwieback, S., Crow, W., Dorigo, W., and Wagner, W.: Recent advances in (soil moisture) triple collocation analysis, International Journal of Applied Earth Observation and Geoinformation, 45, 200–211, https://doi.org/10.1016/j.jag.2015.09.002, 2016.

He, X., Li, Y., Liu, S., Xu, T., Chen, F., Li, Z., Zhang, Z., Liu, R., Song, L., Xu, Z., Peng, Z., and Zheng, C.: Improving regional climate simulations based on a hybrid data assimilation and machine learning method, Hydrology and Earth System Sciences, 27, 1583–1606, https://doi.org/10.5194/hess-27-1583-2023, 2023.

Jung, H. C., Getirana, A., Arsenault, K. R., Kumar, S., and Maigary, I.: Improving surface soil moisture estimates in West Africa through GRACE data assimilation, Journal of Hydrology, 575, 192–201, https://doi.org/10.1016/j.jhydrol.2019.05.042, 2019.

Keller, J. D. and Potthast, R.: AI-based data assimilation: Learning the functional of analysis estimation, https://doi.org/10.48550/ARXIV.2406.00390, 2024.

Khaki, M., Hoteit, I., Kuhn, M., Awange, J., Forootan, E., van Dijk, A. I. J. M., Schumacher, M., and Pattiaratchi, C.: 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, 2017.

Klees, R., Zapreeva, E. A., Winsemius, H. C., and Savenije, H. H. G.: The bias in GRACE estimates of continental water storage variations, Hydrology and Earth System Sciences, 11, 1227–1241, https://doi.org/10.5194/hess-11-1227-2007, 2007.

Kumar, S. V., Zaitchik, B. F., Peters-Lidard, C. D., Rodell, M., Reichle, R., Li, B., Jasinski, M., Mocko, D., Getirana, A., De Lannoy, G., Cosh, M. H., Hain, C. R., Anderson, M., Arsenault, K. R., Xia, Y., and Ek, M.: Assimilation of Gridded GRACE Terrestrial Water Storage Estimates in the North American Land Data Assimilation System, Journal of Hydrometeorology, 17, 1951–1972, https://doi.org/10.1175/JHM-D-15-0157.1, 2016.

Kumar, S. V., Dirmeyer, P. A., Peters-Lidard, C. D., Bindlish, R., and Bolten, J.: Information theoretic evaluation of satellite soil moisture retrievals, Remote Sensing of Environment, 204, 392–400, https://doi.org/10.1016/j.rse.2017.10.016, 2018.

Landerer, F. W. and Swenson, S. C.: Accuracy of scaled GRACE terrestrial water storage estimates, Water Resources Research, 48, https://doi.org/10.1029/2011WR011453, 2012.

Landerer, F. W., Flechtner, F. M., Save, H., Webb, F. H., Bandikova, T., Bertiger, W. I., Bettadpur, S. V., Byun, S. H., Dahle, C., Dobslaw, H., Fahnestock, E., Harvey, N., Kang, Z., Kruizinga, G. L. H., Loomis, B. D., McCullough, C., Murböck, M., Nagel, P., Paik, M., Pie, N., Poole, S., Strekalov, D., Tamisiea, M. E., Wang, F., Watkins, M. M., Wen, H.-Y., Wiese, D. N., and Yuan, D.-N.: Extending the Global Mass Change Data Record: GRACE Follow-On Instrument and Science Data Performance, Geophysical Research Letters, 47, https://doi.org/10.1029/2020GL088306, 2020.

Lecomte, H., Rosat, S., and Mandea, M.: Gap filling between GRACE and GRACE-FO missions: assessment of interpolation techniques, Journal of Geodesy, 98, 107, https://doi.org/10.1007/s00190-024-01917-3, 2024.

Lenczuk, A., Weigelt, M., Kosek, W., and Mikocki, J.: Autoregressive Reconstruction of Total Water Storage within GRACE and GRACE Follow-On Gap Period, Energies, 15, 4827, https://doi.org/10.3390/en15134827, 2022.

Libero, G. and Ciriello, V.: Dominant spatiotemporal structures in total water storage anomalies, Advances in Water Resources, 203, 105 015, https://doi.org/https://doi.org/10.1016/j.advwatres.2025.105015, 2025.

Liu, D., Mishra, A. K., Yu, Z., Lü, H., and Li, Y.: Support vector machine and data assimilation framework for Groundwater Level Forecasting using GRACE satellite data, Journal of Hydrology, 603, 126 929, https://doi.org/10.1016/j.jhydrol.2021.126929, 2021.

Longuevergne, L., Scanlon, B. R., and Wilson, C. R.: GRACE Hydrological estimates for small basins: Evaluating processing approaches on the High Plains Aquifer, USA, Water Resources Research, 46, https://doi.org/https://doi.org/10.1029/2009WR008564, 2010.

Maina, F. Z., Xue, Y., Kumar, S. V., Getirana, A., McLarty, S., Appana, R., Forman, B., Zaitchik, B., Loomis, B., Maggioni, V., and Zhou, Y.: Development of a multidecadal land reanalysis over High Mountain Asia, Scientific Data, 11, 827, https://doi.org/10.1038/s41597-024-03643-z, publisher: Nature Publishing Group, 2024.

Mo, S., Zhong, Y., Forootan, E., Mehrnegar, N., Yin, X., Wu, J., Feng, W., and Shi, X.: Bayesian convolutional neural networks for predicting the terrestrial water storage anomalies during GRACE and GRACE-FO gap, Journal of Hydrology, 604, 127 244, https://doi.org/10.1016/j.jhydrol.2021.127244, 2022.

Nie, W., Zaitchik, B. F., Rodell, M., Kumar, S. V., Arsenault, K. R., Li, B., and Getirana, A.: Assimilating GRACE into a land surface model in the presence of an irrigation-induced groundwater trend, Water Resources Research, 55, 11 274–11 294, 2019.

Reichle, R. H., Koster, R. D., Liu, P., Mahanama, S. P. P., Njoku, E. G., and Owe, M.: Comparison and assimilation of global soil moisture retrievals from the Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E) and the Scanning Multichannel Microwave Radiometer (SMMR), Journal of Geophysical Research: Atmospheres, 112, 2006JD008 033, https://doi.org/10.1029/2006JD008033, 2007.

Retegui-Schiettekatte, L., Schumacher, M., Madsen, H., and Forootan, E.: Assessing daily GRACE Data Assimilation during flood events of the Brahmaputra River Basin, Science of The Total Environment, 975, 179 181, https://doi.org/10.1016/j.scitotenv.2025.179181, 2025.

Rodell, M., Famiglietti, J. S., Wiese, D. N., Reager, J. T., Beaudoing, H. K., Landerer, F. W., and Lo, M.-H.: Emerging trends in global freshwater availability, Nature, 557, 651–659, https://doi.org/10.1038/s41586-018-0123-1, publisher: Nature Publishing Group, 2018.

Ruiz, J. J., Pulido, M., and Miyoshi, T.: Estimating model parameters with ensemble-based data assimilation: A review, Journal of the Meteorological Society of Japan. Ser. II, 91, 79–99, https://doi.org/10.2151/jmsj.2013-201, 2013.

Schumacher, M., Kusche, J., and Döll, P.: A systematic impact assessment of GRACE error correlation on data assimilation in hydrological models, Journal of Geodesy, 90, 537–559, https://doi.org/10.1007/s00190-016-0892-y, 2016.

Shokri, A., Walker, J. P., Van Dijk, A. I. J. M., and Pauwels, V. R. N.: Performance of Different Ensemble Kalman Filter Structures to Assimilate GRACE Terrestrial Water Storage Estimates Into a High-Resolution Hydrological Model: A Synthetic Study, Water Resources Research, 54, 8931–8951, https://doi.org/10.1029/2018WR022785, 2018.

Soltani, S. S., Ataie-Ashtiani, B., and Simmons, C. T.: Review of assimilating GRACE terrestrial water storage data into hydrological models: Advances, challenges and opportunities, Earth-Science Reviews, 213, 103 487, https://doi.org/10.1016/j.earscirev.2020.103487, 2021.

Soltani, S. S., Ataie-Ashtiani, B., Al Bitar, A., Simmons, C. T., Younes, A., and Fahs, M.: Assimilating multivariate remote sensing data into a fully coupled subsurface-land surface hydrological model, Journal of Hydrology, 641, 131 812, https://doi.org/10.1016/j.jhydrol.2024.131812, 2024.

Springer, A., Karegar, M. A., Kusche, J., Keune, J., Kurtz, W., and Kollet, S.: Evidence of daily hydrological loading in GPS time series over Europe, Journal of Geodesy, 93, 2145–2153, https://doi.org/10.1007/s00190-019-01295-1, 2019.

Stoffelen, A.: Toward the true near-surface wind speed: Error modeling and calibration using triple collocation, Journal of Geophysical Research: Oceans, 103, 7755–7766, https://doi.org/10.1029/97JC03180, 1998.

Su, H., Yang, Z., Dickinson, R. E., Wilson, C. R., and Niu, G.: Multisensor snow data assimilation at the continental scale: The value of Gravity Recovery and Climate Experiment terrestrial water storage information, Journal of Geophysical Research: Atmospheres, 115, 2009JD013 035, https://doi.org/10.1029/2009JD013035, 2010.

Tang, G., Clark, M. P., Papalexiou, S. M., Ma, Z., and Hong, Y.: Have satellite precipitation products improved over last two decades? A comprehensive comparison of GPM IMERG with nine satellite and reanalysis datasets, Remote Sensing of Environment, 240, 111 697, https://doi.org/10.1016/j.rse.2020.111697, 2020.

Tangdamrongsub, N., Steele-Dunne, S. C., Gunter, B. C., Ditmar, P. G., and Weerts, A. H.: Data assimilation of GRACE terrestrial water storage estimates into a regional hydrological model of the Rhine River basin, Hydrology and Earth System Sciences, 19, 2079–2100, https://doi.org/10.5194/hess-19-2079-2015, 2015.

Tangdamrongsub, N., Steele-Dunne, S. C., Gunter, B. C., Ditmar, P. G., Sutanudjaja, E. H., Sun, Y., Xia, T., and Wang, Z.: Improving estimates of water resources in a semi-arid region by assimilating GRACE data into the PCR-GLOBWB hydrological model, Hydrology and Earth System Sciences, 21, 2053–2074, https://doi.org/10.5194/hess-21-2053-2017, 2017.

Tapley, B. D., Bettadpur, S., Watkins, M., and Reigber, C.: The gravity recovery and climate experiment: Mission overview and early results, Geophysical Research Letters, 31, https://doi.org/10.1029/2004GL019779, 2004.

Tian, S., Tregoning, P., Renzullo, L. J., Van Dijk, A. I., Walker, J. P., Pauwels, V. R., and Allgeyer, S.: Improved water balance component estimates through joint assimilation of GRACE water storage and SMOS soil moisture retrievals, Water Resources Research, 53, 1820–1840, 2017.

Uz, M., Atman, K. G., Akyilmaz, O., Shum, C., Keleş, M., Ay, T., Tandoğdu, B., Zhang, Y., and Mercan, H.: Bridging the gap between GRACE and GRACE-FO missions with deep learning aided water storage simulations, Science of The Total Environment, 830, 154 701, https://doi.org/10.1016/j.scitotenv.2022.154701, 2022.

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

Van Loon, S. and Fletcher, S. J.: Foundations for Universal Non-Gaussian Data Assimilation, Geophysical Research Letters, 50, e2023GL105 148, https://doi.org/https://doi.org/10.1029/2023GL105148, e2023GL105148 2023GL105148, 2023.

Vishwakarma, B. D., Horwath, M., Devaraju, B., Groh, A., and Sneeuw, N.: A Data-Driven Approach for Repairing the Hydrological Catchment Signal Damage Due to Filtering of GRACE Products, Water Resources Research, 53, 9824–9844, https://doi.org/https://doi.org/10.1002/2017WR021150, 2017.

Wahr, J., Molenaar, M., and Bryan, F.: Time variability of the Earth's gravity field: Hydrological and oceanic effects and their possible detection using GRACE, Journal of Geophysical Research: Solid Earth, 103, 30 205–30 229, https://doi.org/10.1029/98JB02844, 1998.

Whitaker, J. S. and Hamill, T. M.: Evaluating methods to account for system errors in ensemble data assimilation, Monthly Weather Review, 140, 3078–3089, https://doi.org/10.1175/MWR-D-11-00276.1, 2012.

Wongchuig, S., Paiva, R., Siqueira, V., Papa, F., Fleischmann, A., Biancamaria, S., Paris, A., Parrens, M., and Al Bitar, A.: Multi-Satellite Data Assimilation for Large-Scale Hydrological-Hydrodynamic Prediction: Proof of Concept in the Amazon Basin, Water Resources Research, 60, e2024WR037 155, https://doi.org/10.1029/2024WR037155, 2024.

Zhang, X., Li, J., Dong, Q., Wang, Z., Zhang, H., and Liu, X.: Bridging the gap between GRACE and GRACE-FO using a hydrological model, Science of The Total Environment, 822, 153 659, https://doi.org/10.1016/j.scitotenv.2022.153659, 2022.

Zhao, K. and Li, X.: Estimating terrestrial water storage changes in the Tarim River Basin using GRACE data, Geophysical Journal International, 211, 1449–1460, 2017.

Zhao, L. and Yang, Z.-L.: Multi-sensor land data assimilation: Toward a robust global soil moisture and snow estimation, Remote Sensing of Environment, 216, 13–27, https://doi.org/10.1016/j.rse.2018.06.033, 2018.