Hydrol. Earth Syst. Sci., 30, 4969–4983, 2026
https://doi.org/10.5194/hess-30-4969-2026

Hydrology and Earth System Sciences
Open Access

Beyond GRACE: Evaluating the benefits of NGGM and MAGIC for precipitation estimation over Europe

Muhammad Usman Liaqat[1], **Stefania Camici**[1], **Francesco Leopardi**[1,2], **Jaime Gaona**[1], and **Luca Brocca**[1]

[1]Institute of Geosciences, National Research Council, Via Madonna Alta, 126, 06128 Perugia, Italy
[2]Department of Civil and Environmental Engineering, University of Perugia, 06125 Perugia, Italy

Correspondence: Muhammad Usman Liaqat (muhammadusman.liaqat@cnr.it)

Received: 30 July 2025 – Discussion started: 17 October 2025
Revised: 27 July 2026 – Accepted: 28 July 2026 – Published: 7 August 2026

Abstract. The Gravity Recovery and Climate Experiment (GRACE) and its Follow-On (GRACE-FO) missions provide observations of terrestrial water storage (TWS) dynamics from regional to global scales. However, their coarse spatio-temporal resolution limits the applicability in many hydrological studies. A joint collaboration between the National Aeronautics and Space Administration (NASA) and the European Space Agency (ESA), known as the Mass Change And Geosciences International Constellation (MAGIC), aims to launch a new paired mission, i.e., GRACE-C and NGGM (Next Generation Gravity Mission), to improve the monitoring of extreme events such as floods and droughts.

The primary objective of this study is to examine the impact of the expected improvement in the spatial-temporal resolution and accuracy of NGGM and MAGIC on precipitation estimation by developing multiple synthetic experiments over Europe. The study employed the well-known SM2RAIN algorithm to estimate the precipitation accumulated between two consecutive TWS measurements. The total amount of water in the soil from the fifth generation ECMWF reanalysis for the land (ERA5L) is used as a proxy of TWS over the period 2003–2012, while ERA5L precipitation serves as a reference dataset. The results showed that SM2RAIN exhibited satisfactory performance at a daily temporal resolution, with a mean correlation coefficient of 0.86. Good agreement was obtained across most of Europe except in some areas of northern Italy, northeastern Europe (Estonia, Latvia) and some coastal regions. Synthetic experiments were developed by degrading the temporal resolution (5–10, 15, and 30 d) and by adding Gaussian errors to TWS data ranging from 1 to 40 mm. The results showed robust performance under moderate temporal degradation (5–10 d) and low measurement errors (< 5 mm), with correlations generally exceeding 0.75. However, performance deteriorates significantly when errors exceed 10–20 mm, with correlations dropping below 0.40 for an error of 42 mm. These findings identify critical design thresholds for NGGM/MAGIC missions, demonstrating that achieving a measurement accuracy better than 5–10 mm is crucial for reliable precipitation estimation across diverse European hydroclimatic conditions.

1 Introduction

Precipitation is one of the primary components of the hydrological cycle, the precise estimation of spatio-temporal distribution of precipitation plays a vital role in assessing extreme events (Vrochidou et al., 2013), water resources management (Dorigo et al., 2021), ecological assessments (Yu et al., 2024) and hydrological research (McMillan et al., 2011; Meles et al., 2024).

Primary methods to estimate precipitation include ground observations (rain gauges and radar), reanalysis datasets and satellite based products (Song et al., 2020; Yuan et al., 2020). Although rain gauges are marked as one of the reliable sources of precipitation measurement, their spatial coverage, especially in mountainous and data-scarce regions, remains a primary limitation for their applicability on a global scale (Ahrens, 2006; Herrnegger et al., 2018; Kidd et al., 2017). Meteorological radar provides high-resolution spatial estimates of precipitation. However, it is constrained by limited accessibility and high operational costs (Sun et al., 2018). Reanalysis datasets such as the 5th reanalysis from

Published by Copernicus Publications on behalf of the European Geosciences Union.

the European Centre for Medium-Range Weather Forecasting (ECMWF), ERA5, have been developed and widely used from regional to global scales. Nevertheless, the accuracy of the reanalysis product depends upon the multi-source data, including in-situ observations and satellite products, which constrains its applicability in a data scarce mountainous environment (Ansari et al., 2022; Li et al., 2020; Liaqat et al., 2022). Satellite based precipitation products with their high spatial and temporal coverage, provide an alternative solution for precipitation estimation.

There are two primary methods for precipitation estimation from satellite, including top-down and bottom-up approaches (Brocca et al., 2014). In the top-down approach, precipitation is based on the emitted or reflected radiation from clouds or rain droplets derived from a combination of Geostationary (GEO) and Low Earth Orbiting (LEO) satellites sensors (Levizzani et al., 2020). In the bottom-up approach, precipitation is computed by surface soil moisture (SM) observations inverting the soil water balance in the so-called SM2RAIN algorithm (Brocca et al., 2013, 2014). A key difference between two approaches is that the top-down provides instantaneous measurements of precipitation which can be underestimated when the satellites do not pass over the precipitation event while the bottom-up approach gives accumulated precipitation between two subsequent SM measurements.

SM2RAIN has already been implemented from regional scale in South Asia (Lai et al., 2022) to global scale (Brocca et al., 2019), in Africa (Hengl et al., 2021; Islam, 2022), America (Paredes-Trejo et al., 2019; Satgé et al., 2020), Australia (Chua et al., 2022) and Europe (Moges et al., 2022). Various SM2RAIN products have been developed by forcing different SM satellite observations into the algorithm, including ASCAT (Advance Scatterometer (Brocca et al., 2019), Soil Moisture and Ocean Salinity (SMOS, Brocca et al., 2016), Advanced Microwave Scanning Radiometer 2 (AMSR2, Tarpanelli et al., 2017), Sentinel-1 (Filippucci et al., 2022).

In this study, SM is replaced for the first time with terrestrial water storage (TWS) to estimate precipitation between two consecutive measurements. SM2RAIN based on surface SM relies on top few centimeters of soil. This layer is quite sensitive to dense vegetation, frozen soils, and snow-covered locations, and it saturates quickly after intense precipitation. Gravimetry, on the other hand, observes TWS integrating water variations over the full soil column and shallow subsurface. This makes it less sensitive to surface saturation and vegetation impacts; however, the detection of precipitation signals in gravimetric observations is directly constrained by measurement uncertainty. As precipitation represents the primary source of TWS, its variations inherently reflect cumulative water inputs and losses at larger spatial scales (Zhong et al., 2025). These characteristics make TWS a suitable choice, potentially better than SM, for estimating accumulated precipitation at sub monthly scales. However, this potential has remained largely unexplored because TWS estimates derived from previous gravity missions, including the Gravity Recovery and Climate Experiment (GRACE, Tapley et al., 2004), Gravity Field and steady-state Ocean Circulation Explorer (GOCE, Drinkwater et al., 2003), and GRACE Follow-On (GRACE-FO, Landerer et al., 2020), are characterized by coarse spatial (∼400 km) and, importantly, temporal (monthly) resolutions, which limit their ability to accurately capture rainfall variability, except in regions experiencing intense precipitation, such as the tropics. Currently, daily-scale TWS estimates can only be obtained through the integration of gravimetric observations with hydrological or land surface models (Croteau et al., 2020). Thus, it is indispensable to work on high temporal resolution satellite-based solutions from daily to weekly scales to improve our capability to estimate global scale precipitation.

A joint collaboration between the National Aeronautics and Space Administration (NASA) and the European Space Agency (ESA), known as the Mass change And Geosciences International Constellation (MAGIC, Daras et al., 2024), was initiated to advance gravity-based monitoring of mass transport within the Earth system. The ultimate objective of this collaboration is to enhance current observing capabilities through next-generation high-resolution gravity missions, thereby improving the monitoring of extreme events such as natural hazards, droughts, and floods. The European Next-Generation Gravity Mission (NGGM, Daras et al., 2024; Haagmans and Tsaoussi, 2020), a component of MAGIC, is presently undergoing Phase A Extension activities as the first mission opportunity under ESA's Future EO Program. Coordinated with the NASA-led GRACE-C component, NGGM is designed to drastically improve both the temporal sampling and effective spatial resolution of TWS observations compared with the GRACE and GRACE-FO missions. Current mission studies indicate a target temporal sampling on the order of sub-weekly days and an effective spatial resolution approaching approximately 100 km at regional scales, for equivalent water height accuracies of approximately 5–10 mm under target performance conditions (Daras, 2023). Combined with further improvements in orbit and attitude modelling, these advances are expected to enable the generation of fast track gravity products on sub-weekly time scales, reducing temporal aliasing while enhancing signal to-noise characteristics.

Based on that, the objectives of this study are: (1) to examine the performance of SM2RAIN algorithm to estimate precipitation using TWS data as input, (2) to investigate the impact of the improved accuracy and spatiotemporal resolution of NGGM and MAGIC on precipitation by using the SM2RAIN algorithm on a European scale, and (3) to analyse the impact of different synthetic mission configurations by adding Gaussian error into simulated TWS observations used as input to the SM2RAIN algorithm. It should be noted that the purpose of this study is not to perform an operational validation of precipitation estimates, but rather to conduct synthetic experiments to assess the feasibility of retrieving precipitation from TWS and evaluating the sensitivity of the SM2RAIN algorithm to the temporal sampling and accuracy expected from future NGGM and MAGIC gravity missions.

2 Material and methods

2.1 Study area

The study focuses on the European region, specifically the area with latitudes 30° to 60° N and longitudes 10° W to 30° E (see Fig. 1). This extensive domain is characterized by diverse climatic zones, each playing a critical role in shaping precipitation dynamics and hydrological processes. Specifically, according to the Köppen-Geiger climate classification (Beck et al., 2023), the European region exhibits significant climatic diversity: the Oceanic Climate (Cfb), mainly evident in Western Europe along the Atlantic coastline and largely regulated by the proximity to the prevailing westerlies and ocean. It is characterized by mild air temperatures and relatively consistent high precipitation throughout the year. In contrast, the Continental Climate (Dfb, Dfc), largely prevails in Central and Eastern Europe where seasonal extremes, especially for air temperature, are more pronounced, characterised by cold winters and warm summers. Precipitation is generally moderate with peak events in summer. The Mediterranean Climate (Csa, Csb) is distinguished by hot dry summers and mild wet winters which largely dominate in Southern Europe and in the Mediterranean Basin. Further, the Polar and Subarctic Climates (Dfc, ET) are experienced in Northern Europe including Scandinavia, and high-altitude regions such as the Alps featured by long harsh winters and short summers. Precipitation is generally low but persistent, with significant amounts of snowfall during the winter, feeding into glacial and snowmelt-driven hydrological processes.

2.2 ERA5L Soil Moisture and Precipitation data

The study used volumetric soil moisture content and total precipitation data from the ERA5-Land, the land component of the global reanalysis product ERA5 developed by the European Centre for Medium-Range Weather Forecasts (Muñoz-Sabater et al., 2021). In particular, the volumetric soil moisture content (m^3 / m^3) at four soil depth layers (i.e., Layer 1, 0–7 cm, Layer 2, 7–28 cm, Layer 3, 28–100 cm, Layer 4, 100–289 cm) were utilized to estimate TWS. Since ERA5L does not explicitly represent groundwater storage, the TWS is derived by computing, for each layer, the product of soil moisture content and layer thickness, which converts volumetric water content into water storage depth within the soil column. The final TWS estimate, in mm, can be obtained by adding the water stored in all four layers. The equation used to calculate TWS is:

$$\mathrm{TWS} = (\mathrm{SM}_1 \cdot 0.07) + (\mathrm{SM}_2 \cdot 0.21) + (\mathrm{SM}_3 \cdot 0.72) + (\mathrm{SM}_4 \cdot 1.89) \cdot 1000 \quad (1)$$

where SM_i depicts volumetric soil moisture content in (m^3 / m^3) for the i-th layer. The constants (0.07, 0.21, 0.72, and 1.89) are the thicknesses in meters of the corresponding soil layers. The calculation provides the TWS in the soil column in meters of water equivalent. Although this ERA5L TWS proxy representation does not explicitly include additional storage components such as snow water equivalent, groundwater, surface water, or canopy interception – components included in the total signal observed by GRACE and GRACE-FO- it shows a strong correlation with GRACE based TWS (Fig. S1 in the Supplement). This agreement can be explained by the fact that, over the selected European domain, the omission of these components is expected to have a limited impact on the present sensitivity experiments. In particular, canopy interception and surface water storage are generally negligible at the ~100 km scale, while snow water equivalent may locally influence storage variability in northern and mountainous regions but has a limited overall contribution during the study period.

ERA5L total precipitation was considered as a benchmark for performance assessment of SM2RAIN algorithm. ERA5L is available at 0.1° and hourly temporal resolution from 1950 to present. In this study, the analysis period was constrained to the timeframe between 2003 and 2012, and both TWS and precipitation were resampled on a common 100 km grid using bilinear interpolation at a daily scale. This smooths out small scale spatial variability, particularly in mountainous regions like the Alps or the Scandinavian mountains. Note that this study is not designed to reproduce the full spatial resolution of the GRACE and GRACE-FO observations but instead explores the sensitivity of the precipitation retrieval to temporal sampling and measurement uncertainty. To support this choice, we conducted a set of scale consistency and sensitivity analyses (Figs. S1 and S2). The results demonstrate that resampling to a common analysis grid preserves gravity relevant information and does not introduce artificial small scale signals, thereby supporting the methodological approach adopted in the synthetic experiments.

2.3 The SM2RAIN algorithm

A modified version of the SM2RAIN algorithm, i.e., using TWS instead of soil moisture, was applied to estimate precipitation between two successive TWS measurements in the time interval dt, by inverting soil water balance equation as follows:

$$Z^* \frac{\mathrm{dTWS}^*(t)}{\mathrm{d}t} = p(t) - r(t) - e(t) - g(t) \quad (2)$$

Here Z^* [mm] represents the water capacity of the soil layer, calculated as the product between the depth of the soil layer

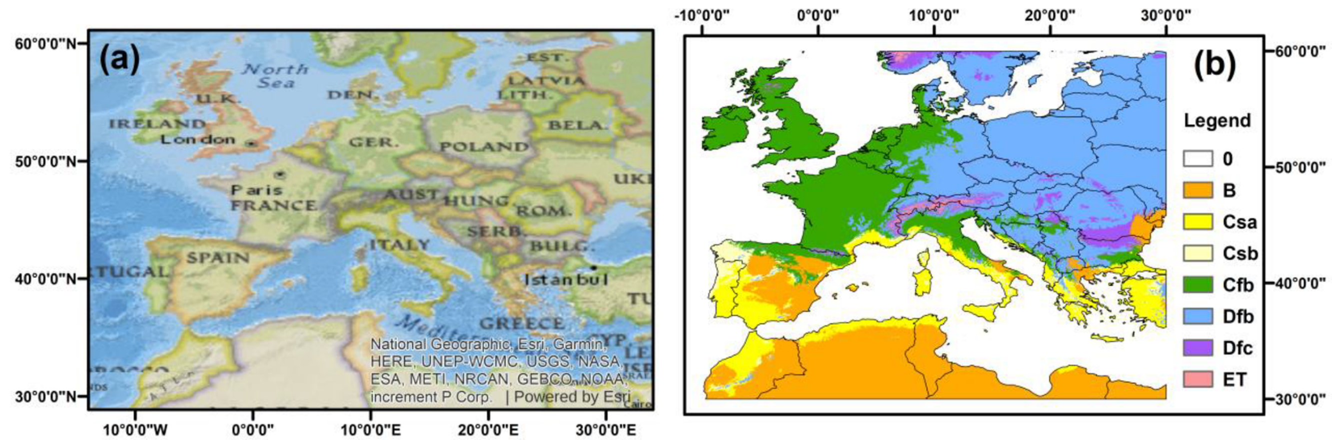

Figure 1. (a) Study domain covering the European region used in this analysis. The spatial grid corresponds to the ERA5L dataset resampled to approximately 100 km spatial resolution. The analysis period spans 2003–2012, corresponding to 3653 daily time steps used in the SM2RAIN experiments. The base map is derived from OpenStreetMap data © OpenStreetMap contributors (available under the Open Database License). **(b)** Köppen–Geiger climate classification across the study domain used to assess the influence of climatic conditions on the performance of precipitation estimation from terrestrial water storage signals. The major climate classes include arid (B), Mediterranean climates (Csa, Csb), temperate oceanic climates (Cfb), cold continental climates (Dfb, Dfc), and tundra climates (ET).

and its porosity. The term TWS* refers to the relative TWS [–] (i.e., ranging between 0 and 1); t is the time, p is the precipitation rate [mm t^{-1}], $r(t)$ is the surface runoff rate [mm t^{-1}], $e(t)$ is the evaporation rate [mm t^{-1}] and $g(t)$ is the drainage rate [mm t^{-1}]. During precipitation events, evaporation and surface runoff are assumed to be negligible (Brocca et al., 2015). The drainage rate is dependent on TWS, which is derived from Famiglietti and Wood's (Famiglietti and Wood, 1994). Thus, the precipitation rate can be estimated by simplifying Eq. (2) as follows:

$$p(t) = Z^* \frac{\mathrm{dTWS}^*(t)}{\mathrm{d}t} + as(t)^b \quad (3)$$

Where a (mm t^{-1}) and b (–) refer to the saturated hydraulic conductivity and the exponent in the (Famiglietti and Wood, 1994) equation, respectively. Calibration is performed independently at each grid cell (point-by-point) using a gradient-based optimization function. The procedure estimates five parameters: Z^*, a, b, T_{pot}, and T_{base}. The estimation of parameters a, b, and Z^* is needed to compute the precipitation between two consecutive TWS* measurements. The filter parameters (T_{pot}, and T_{base}) have narrow interquartile ranges, indicating consistent noise reduction across grid cells. Before running the algorithm, the noise in the TWS signal is reduced by applying the exponential filter as described by Brocca et al. (2019), where the characteristic time length parameter (T) is modeled as a function of soil moisture using a two-parameter power law relationship. The values of the two parameters (T_{pot}, and T_{base}) of the modified exponential filter equation, and the a, b, and Z^* parameter values are estimated through calibration against the ERA5L precipitation dataset. The parameter bounds are explicitly defined in the code as follows: Z^* ranges between 10 and 1200, a between 0 and 160, b between 1 and 50, T_{pot} between 0.05 and 0.75, and T_{base} between 0.05 and 3.0. Optimization is carried out using MATLAB's fmincon function with the active set algorithm, allowing a maximum of 300 iterations and 500 function evaluations. The calibration is carried out by minimizing an objective function based on the Nash–Sutcliffe efficiency (NS, Nash and Sutcliffe, 1970), computed between simulated and reference precipitation (e.g., ERA5L).

To evaluate the spatial consistency and the robustness of the calibration, we computed summary statistics of the calibrated parameters across the European domain. These values correspond to unrealistic parameter values across other parameterizations, e.g., mountainous or cold regions. The distribution of calibrated parameters (Table S1) supports the occurrence of substantial parameter variability and equifinality within this calibration framework. The storage capacity parameter Z^* ranges from 38.85 to 1008.70, with a median value of 506.80 and an interquartile range (IQR) of 432.67–584.10. These values are substantially larger than those typically reported for surface soil moisture based SM2RAIN applications, reflecting the deeper and more integrated nature of terrestrial water storage signals derived from gravimetry. The detail information about these calibrated parameters can be found in Table S1.

Regarding the validation strategy, the current implementation uses the full available time series at each grid cell for both calibration and performance assessment. An explicit temporal train/validation split or spatial cross validation was not applied. This design choice is consistent with the primary objective of the study, which is to conduct a controlled synthetic sensitivity assessment of precipitation retrieval performance under different NGGM/MAGIC sampling and error scenarios, rather than to provide operational out of sample validation of a precipitation product.

2.4 Setup of the synthetic experiments

The study employed the SM2RAIN algorithm to estimate the precipitation accumulated between two consecutive TWS measurements (Brocca et al., 2014). The study considered three types of experiments: (i) feasibility experiments, in which ERA5L TWS proxy were used as a proxy for perfect storage retrievals and to evaluate the capability of SM2RAIN to reconstruct precipitation from storage variations; (ii) synthetic experiments simulating the expected sampling and error characteristics of NGGM/MAGIC missions, (iii) reference scenarios, which represent GRACE/GRACE-FO-like temporal sampling conditions.

As a first step, the ERA5L TWS proxy at daily time scale was used as input to SM2RAIN, while ERA5L precipitation was used as reference. This step was carried out to check the reliability of SM2RAIN in estimating precipitation from TWS. The simulated precipitation was aggregated to 15 and 30 d scales for comparative analysis and clearer visualization of performance. The choice to use ERA5L as a self consistent "proxy truth" follows established methodology used in gravity mission design and Phase A simulation studies for GRACE, GRACE-FO, MAGIC, and NGGM (e.g., Tapley et al., 2004; Landerer et al., 2020; Haagmans and Tsaoussi, 2020; Daras et al., 2023). Synthetic mission assessments typically rely on a single physically coherent dataset (model or reanalysis) to ensure compatibility between water balance terms and to isolate the effects of instrument accuracy and sampling strategy. Accordingly, ERA5L was used not for operational validation but to provide a controlled environment for evaluating the relative performance of different NGGM/-MAGIC configurations.

In a second step, synthetic experiments were conducted by degrading daily TWS time series to 5 d intervals, representing the temporal sampling anticipated for future NGGM/MAGIC missions. Temporal resampling is implemented using a moving average aggregation scheme, followed by aggregation to 15 and 30 d scales for clearer visualization of performance. A fixed window (e.g., five samples for the 5 d scenario) is applied to both the TWS signal and its time axis. The smoothed values are then embedded between the first and last original observations to preserve endpoints, and the resulting series is linearly interpolated back to the original daily time grid. This ensures temporal alignment while emulating reduced mission sampling frequency. This design choice makes sure that the experiments are consistent with each other and lets the analysis separate the effects of temporal sampling and measurement uncertainty in the TWS signal. The exponential filter described by Brocca et al. (2019) is subsequently applied within the SM2RAIN framework to the resampled (and noise-perturbed) TWS signal. As reference, the simulated precipitation generated from the first step was used. Although systematic errors from the initial step may carry over to the subsequent step, the purpose of the synthetic experiments is to assess the comparative sensitivity of precipitation retrieval to NGGM/MAGIC sampling configurations, rather than to provide an independent validation.

Further, Gaussian errors (1.9, 4.2, 19 and 42 mm) were incorporated into the resampled TWS data to conduct a first-order sensitivity analysis of precipitation retrieval under measurement uncertainty representative of the expected NGGM/MAGIC error magnitudes (Daras et al., 2023). The Gaussian noise we introduce is a simplified error model. In reality, gravimetric mission errors may have spatial and temporal correlations, depend on orbital geometry of the mission, and vary regionally (e.g., with topography), but the purpose of this noise model is to assess the sensitivity of the precipitation retrieval to measurement uncertainty. Thus, the Gaussian assumption must be understood as a linear approximation of the measurement uncertainty rather than the complete expected error structure of future gravity missions. Noise is generated using a normal distribution and is assumed to be temporally and spatially uncorrelated, providing a controlled baseline for sensitivity analysis (Landerer et al., 2020; Daras et al., 2023).

The exponential filter described by Brocca et al. (2019) is subsequently used within the SM2RAIN framework to the resampled and noise-perturbed TWS signal prior to precipitation inversion. To provide a reference comparable to the capability of current satellite gravimetry missions, a GRACE-FO-like synthetic configuration was defined using approximately ∼25 mm equivalent water height uncertainty and 30 d temporal sampling. This uncertainty level is within the typical gravimetric error levels described in the ESA NGGM Mission Requirements Document (Daras, 2023). This configuration does not use actual GRACE observations but instead mimics present-day gravimetry performance by degrading the ERA5L TWS proxy signal. These error values are the expected uncertainty values from the NGGM mission for the short-term time scales and for spatial resolutions of 400 and 800 km (see Table 11 in Daras et al., 2023).

In a third step, the combined effects of temporal aggregation and error level were assessed by considering temporal intervals of 5, 10, 15, and 30 d for TWS measurements and error levels ranging from 0 to 40 mm. The results were evaluated over 100 randomly selected grid points across Europe. The final step was carried out to investigate the potential improvements of NGGM and MAGIC missions with respect to GRACE\-FO mission. The obtained results were tested with a temporal aggregation of 15 and 30 d, as having TWS measurements every 5 d and differentiating the signal, precipitation can be estimated reliably for a temporal aggregation 3 times the resolution of the input TWS measurements (Brocca et al., 2013).

For each experiment configuration model performance was evaluated using complementary metrics including bias (mm), correlation coefficient (R, –), and root mean square error (RMSE, mm). This approach is consistent with previous studies (e.g., Knoben et al., 2019; Williams, 2025), which recommend reserving skill scores (e.g., Nash Sutcliffe efficiency (NS) index) primarily for model calibration rather than performance evaluation.

3 Results and Discussion

3.1 SM2RAIN reliability using TWS measurements as input

The assessment of the SM2RAIN algorithm for estimating 15 and 30 d precipitation using ERA5L TWS proxy is presented in Fig. 2. The results indicate that SM2RAIN exhibited satisfactory performance, with mean values of R and RMSE equal to 0.86 and 13.31 mm, respectively, for the 15 d and equal to 0.86, 21.06 mm for the 30 d temporal assessment. The mean value of bias indicated a slight underestimation at 15 d (−0.31 mm) with respect to 30 d (−0.23 mm). Significant negative values of bias were observed across most of Europe, with the strongest underestimation noted in central and northeastern regions. Although the Mediterranean and southern Europe also exhibit negative bias, their impact is generally less intense than in northern regions. Based on the statistical analysis, SM2RAIN simulated precipitation shows good agreement across most of Europe, with exceptions in the Alpine areas in northern Italy, in the northeastern states (Estonia, Latvia) and in the coastal regions of Norway. The lower performance in these regions can be attributed to several hydrological factors. For instance, the strong orographic precipitation gradients and pronounced subgrid variability of mountainous regions such as the Alps are hard to represent at the 100 km spatial resolution used in this study, leading to inaccuracies in precipitation estimation. In northern latitudes (for example, Estonia, Latvia and coastal Norway), seasonal snow accumulation and melt introduce lags between precipitation input and changes in terrestrial water storage that also weaken the direct relationship assumed by the SM2RAIN inversion. In coastal areas, maritime precipitation variability and groundwater dynamics may further complicate the precipitation storage relationship (Trautmann et al., 2018). These process based explanations are consistent with the climate stratified analysis shown in Tables S2–S3, which shows that model skill is lower in the colder, more continental, snow dominated and continental climate regimes.

The temporal plot of mean simulated and observed precipitation for 15 d aggregation during 2003–2012 further supports the results discussed above, as shown in Fig. 3. The consistent overlap of both low and high precipitation limbs supports the ability of the simulations to capture the observed variability. The performance obtained in this study ($R = 0.86$ for the 15 d aggregation) is comparable to, and in some cases exceeds, the correlations typically reported in previous SM2RAIN applications based on satellite SM observations, where correlations with reference precipitation datasets generally range between 0.6 and 0.8 at regional to global scales (e.g., Brocca et al., 2019). Overall, these results pave the way for assessing the suitability of SM2RAIN to simulate precipitation for the first time using TWS, and they support the development of synthetic experiments for precipitation estimation under future gravity missions.

Additional analyses were conducted to examine the sensitivity of the SM2RAIN based precipitation estimates to regional hydroclimatic conditions and rainfall intensity (Tables S2–S3). The results show that model performance varies across Köppen–Geiger climate classes, with the highest correlations observed in Mediterranean and temperate oceanic climates, where precipitation-driven storage variations dominate the hydrological response. Lower performance is observed in colder continental and high latitude climates, likely due to the influence of snow accumulation, freeze thaw processes, and delayed storage responses.

To further characterize the relationship between the ERA5L TWS proxy used as SM2RAIN input and the ERA5L precipitation as reference, an additional diagnostic analysis was performed. Grid-cell wise temporal correlations were calculated between 15 d (30 d) TWS changes and the corresponding accumulated precipitation as shown in Fig. S4. The results show a generally strong but spatially variable coupling, with spatial median correlation coefficients of 0.69 and 0.58 for the 15 and 30 d respectively. Higher correlations were observed across most of Europe with exceptions in the Alpine areas in northern Italy, in the northeastern states (Estonia, Latvia) and in the coastal regions of Norway, where snow accumulation and melt, evapotranspiration, groundwater exchange, and antecedent storage conditions increasingly influence terrestrial water storage. The lower median correlation at the 30 d aggregation further illustrates that storage progressively integrates multiple hydrological processes over longer timescales rather than responding solely to precipitation. Overall, this analysis confirms that ERA5L precipitation and TWS are physically coupled within the synthetic framework, while also demonstrating that TWS changes cannot be regarded as a direct surrogate for precipitation.

Furthermore, the analysis improved by precipitation intensity indicates that the method performs significantly better for higher intensity precipitation events. This behavior can be explained by the fact that high intensity rainfall events generate larger and more rapid changes in TWS, which are more clearly detectable above background noise and other hydrological processes such as evaporation and drainage. In contrast, low intensity precipitation events often produce weak storage signals that can be masked by measurement of uncertainty and competing fluxes, leading to reduced correlation with the reference dataset. These findings highlight the potential value of next generation gravity missions such as NGGM and MAGIC, whose improved temporal resolution and measurement accuracy are expected to enhance the de-

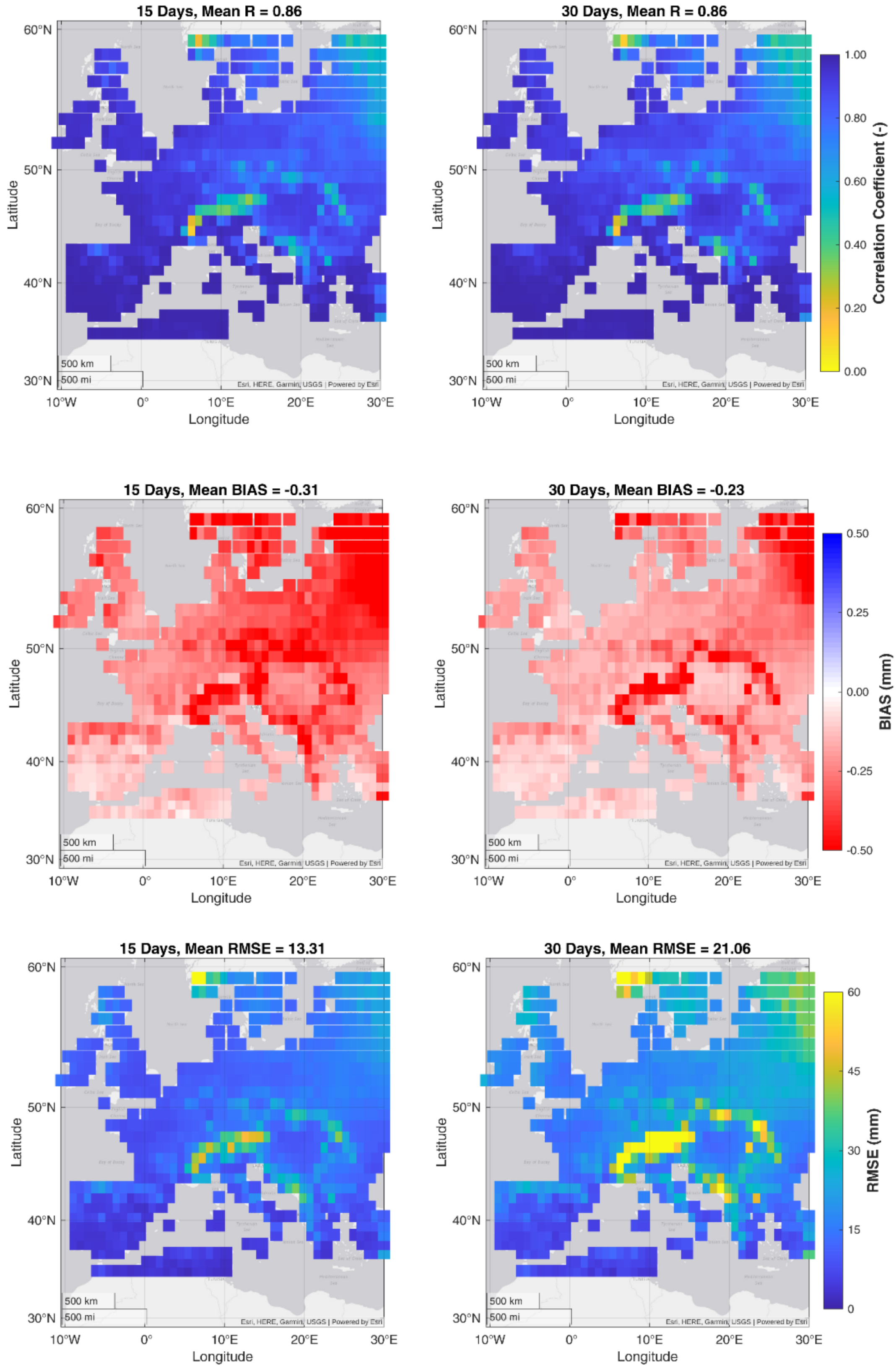

Figure 2. Performances of SM2RAIN algorithm forced with ERA5L TWS proxy against the ERA5L precipitation data over Europe for the period 2003–2012. The two columns refer to 15 and 30 d aggregation period. Rows indicate the correlation coefficient, R (top panel), bias (middle panel) and root mean square error, RMSE (lower panel).

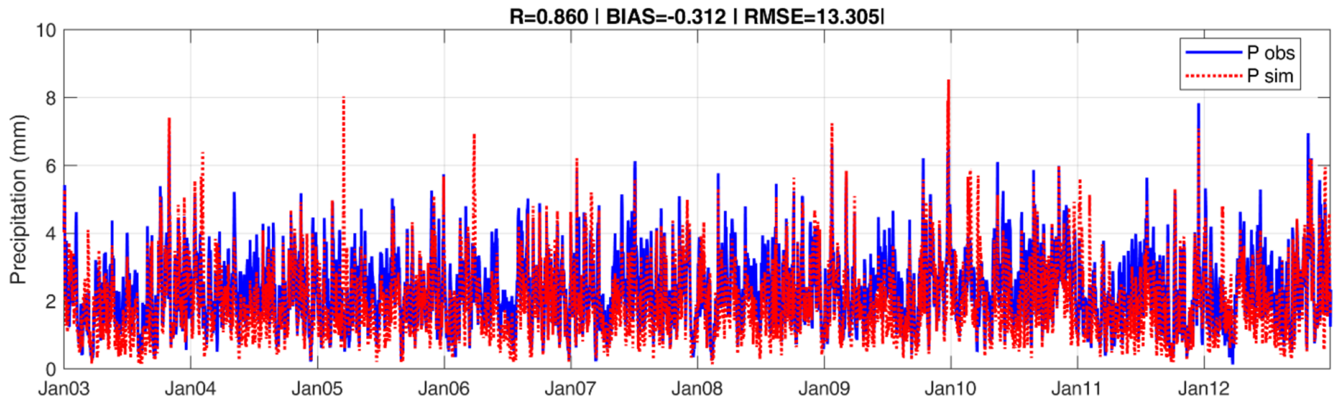

Figure 3. Mean value of precipitation for 15 d aggregation as estimated from TWS through SM2RAIN and obtained from ERA5L precipitation.

tectability of short-term storage changes and improve precipitation estimation, particularly for hydrological extremes.

3.2 Performance of NGGM/MAGIC synthetic TWS for precipitation estimation

In this section, first-order synthetic experiments were designed to assess the performance of the NGGM/MAGIC mission configurations. SM2RAIN derived precipitation during the evaluation stage is used as a reference, while ERA5L TWS proxy aggregated at 5 d interval was used as an input. Measurement errors of 1.9 and 4.2 mm (target error) and 19 and 42 mm (threshold errors) were introduced following Daras et al. (2023).

The temporal resampling of data to a 5 d interval generally maintains a good agreement with the reference precipitation data as evident from all statistical measures with mean R, RMSE, and bias values of 0.95, 9.45, and −0.21 mm, respectively (see Fig. 4, No Error plot). Stronger correlation (R values above 0.88) was observed in the southern Europe and parts of Mediterranean regions. A similar pattern was noted for RMSE values with lower values in southern areas and moderate to higher RMSE values in northern and central Europe. The spatial average of bias exhibited widespread underestimation, with slight overestimation in the Mediterranean.

Introducing a target error from 1.9 to 4.2 mm reduced model's performance as reflected by mean R values equal to 0.91 (−4 %) and 0.87 (−8 %) with respect to the no error configuration, indicating that the reduction is limited when target error is considered. The mean values of RMSE increased more significantly to 10.20 mm (+8 %) and 12.43 mm (+32 %) suggesting that the absolute accuracy begins to depict measurable degradation even with smaller error values while the temporal agreement, measured by R, remains preserved. Spatially, both errors keep higher correlations ($R > 0.80$), consistent error distribution (RMSE < 12 mm) across most of Europe, with slightly less performance in the Eastern European and Scandinavia. Further, the values of bias were found stable with mean values significantly low and equal to −0.01 and 0.04 mm using 1.9 and 4.2 mm error, respectively. The spatial average of bias values indicated a tendency towards underestimation over majority of areas with slightly overestimation in the Mediterranean region. This suggests that low values of errors do not induce systematic algorithmic biases while preserving the optimal model operation to simulate precipitation.

At threshold error (19 to 42 mm), lead to a substantial degradation of model performance with mean R values decreasing to 0.63 (−34 %) and 0.32 (−66 %) and RMSE increasing to 20.99 mm (+121 %) and 24.73 mm (+162 %). The spatial pattern indicates a poor relationship between simulated precipitation and reference precipitation ($R < 0.60$ and RMSE > 20 mm) in several areas (Northern Scandinavia, scattered areas of Eastern Europe, the Mediterranean region, parts of Alps and surrounding mountainous regions). The results demonstrate that larger errors make it extremely difficult for the model to capture and represent meaningful precipitation patterns with 42 mm error level; reasonable performance were kept with 19 mm error level. Finally, the bias analysis illustrates interesting patterns at threshold error. Mean bias varies between 0.15 to 0.04 mm for target and threshold error respectively The spatial inspection demonstrates substantial positive bias under 19 mm error which is evident across southern and central Europe, while mix of both positive and negative bias can be evident under 42 mm indicating more complex patterns, suggesting that extreme errors induce spatially variable systematic distortions rather than uniform bias trends.

The model's performance was also evaluated by aggregating the simulated precipitation for a 30 d period, while keeping the same target and threshold errors, i.e. from 1.9 to 42 mm; results are shown in Table 1. Starting from the baseline without error, the model maintains good R and RMSE values under low error conditions (1.9 and 4.2 mm), At an error level of 4.2 mm, R decreases by up to 7 %, while RMSE

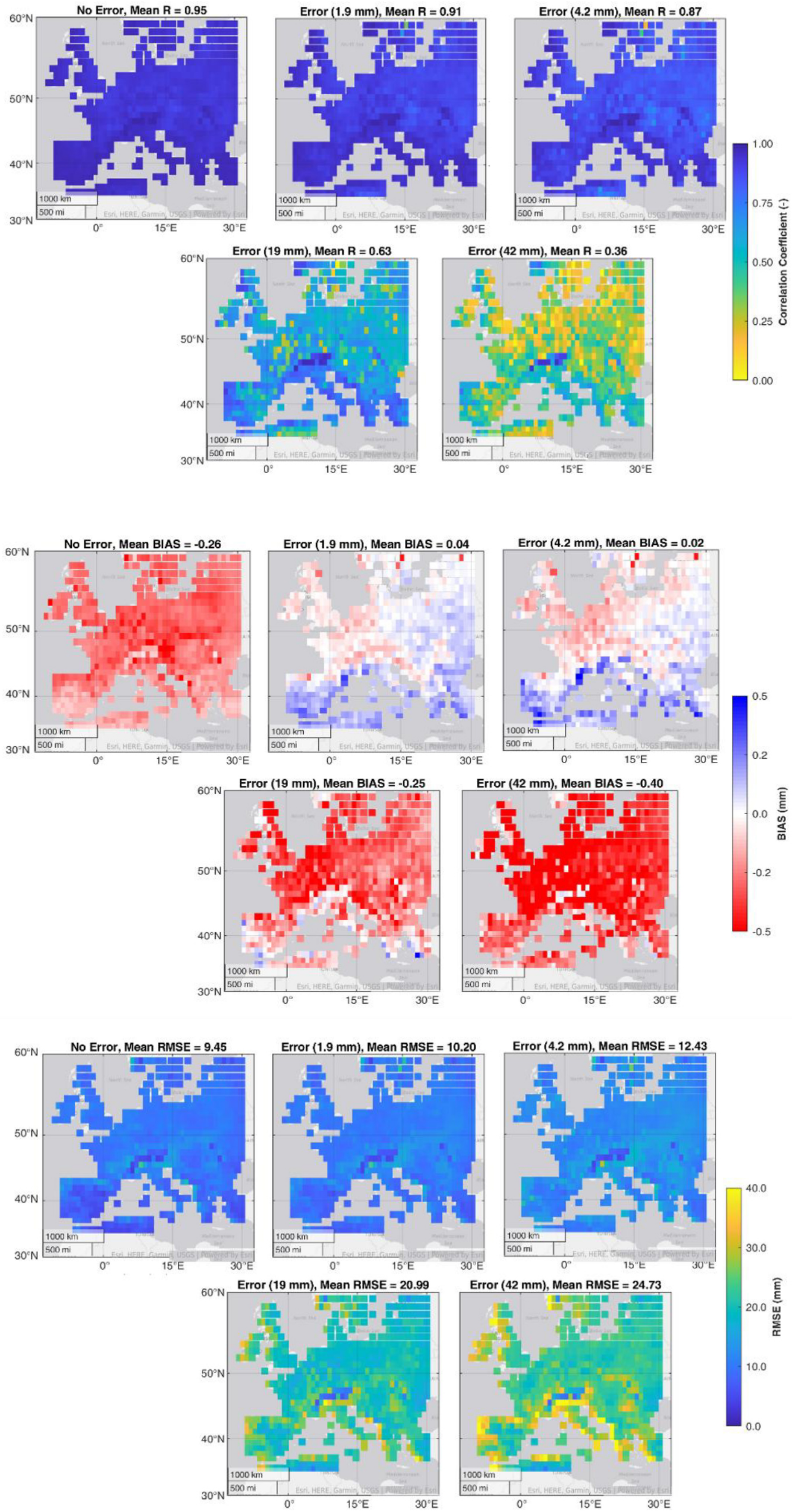

Figure 4. Performance of SM2RAIN algorithm as a proxies of the future NGGM mission forced with ERA5L TWS proxy at 5 d resolution over Europe for the period 2003–2012 and introducing an error from 0 to 42 mm for a 15 d aggregation period. The top panels shows the correlation coefficient, the middle panels the bias, and the lower panels the root mean square error, RMSE.

Table 1. Performance of SM2RAIN algorithm as a proxies of the future NGGM/MAGIC mission forced with ERA5L TWS proxy at 5 d resolution for the period 2003–2012 and introducing an error of 1.9 to 42 mm (top to bottom in rows) for 30 d aggregation period. R, bias and RMSE refer to mean values over Europe.

	R (–)	bias (mm)	RMSE (mm)
No Error	0.95	−0.18	14.44
1.9 mm	0.92	−0.04	14.98
4.2 mm	0.88	0	17.62
19 mm	0.66	0.09	30.92
42 mm	0.32	−0.01	37.17

increases by 22 %. As for 15 d precipitation estimation, the performance deteriorates significantly under large error (19 to 42 mm). It is also noted that model exhibited enhanced correlation stability under lower error range for the 30 d compared to 15 d aggregation period. This suggests that longer temporal aggregation helps to preserve the model's ability to capture precipitation dynamics by smoothing out short-term fluctuations.

Differently from R and RMSE, the bias remains quite stable in the different configurations. This stability is particularly noteworthy when compared to the spatial patterns observed in the 15 d analysis, where regional bias variations were more pronounced. The 30 d aggregation appears to further stabilize the algorithm's systematic performance, likely due to the temporal averaging effect that reduces the impact of short-term measurement inconsistencies.

3.3 Performance of NGGM/MAGIC mission with respect to GRACE\-FO

The impact of varying temporal resolution (5, 10, 15, 30 d) and error level (from 0 to 40 mm) was evaluated at eight randomly selected locations in Italy, Germany, Spain, Greece, Austria, Czechia, Denmark, and Estonia. The results covering the period from 2003 to 2012 for 30 d precipitation estimation are shown in Fig. 5. The findings reveal that the temporal resolution has a significant impact on model performance with varying patterns across error levels. Overall, better temporal resolutions (5–10 d) maintain high R and low RMSE values, even with increasing error, while lower temporal resolutions exhibit sharper drops in correlation as the error grows. However, the rate of degradation in the performance of simulated precipitation varies between countries. Generally, for error values larger than 20 mm, the performance is quite similar for the different temporal resolutions, particularly for the RMSE values. For lower error levels, the differences in the performance for the different temporal resolutions are much larger. This regional representation improves the robustness of the findings by highlighting the model's applicability across diverse climatic conditions.

To obtain more general conclusions regarding the impact of error level and temporal resolution, an additional experiment was conducted using 100 grid cells randomly sampled from all valid land pixels across the European domain (i.e., grid cells with complete TWS and precipitation records and no missing values), as shown in Figs. 6 and S3. To ensure reproducibility of the sampling procedure, the random selection was performed using a fixed random seed. No explicit spatial or climatic stratification was applied in this step, as the experiment was intended to provide a representative domain wide sensitivity assessment rather than a formally stratified evaluation.

The analysis revealed a linear decrease in the R values as a function of error levels and again more stable performance for error levels larger than 20 mm in terms of RMSE. By setting a threshold value for R to 0.7, above which the performance can be considered satisfactory, for 5 d temporal resolution data the results are satisfactory for error levels lower than 15 mm. For 30 d temporal resolution, the results are not satisfactory even for no error conditions. If we consider the expected error level of NGGM\MAGIC to be lower than 10 mm, we can foresee mean performance better than 0.78 (35 mm) in terms of R (RMSE). The error level of GRACE\-FO was found equal to 25 mm and the temporal resolution is 30 d. The performance is equal to 0.40 (55 mm) in terms of R (RMSE). The expected improvements of NGGM\MAGIC with respect to GRACE are therefore highly significant.

Overall, this analysis shows smoother degradation patterns than those observed in individual country analyses, suggesting that regions with more challenging conditions (e.g., Mediterranean areas) are partly compensated when results are averaged across the diverse climatic conditions of Europe. The findings also support the 5–10 d temporal resolution which offers optimal performance for precipitation estimation at a regional scale. The improvement in precipitation estimation with high temporal resolution aligns with previous scientific expectations that shorter revisit times reduce temporal aliasing of storage signals. Similar findings have been reported in previous gravity based hydrological studies (Save et al., 2016; Scanlon et al., 2018), highlighting the significance of sub-monthly sampling for improving hydrological fluxes.

4 Limitations and Future Work

This study conducts a controlled sensitivity experiment utilizing ERA5L as a self-consistent synthetic reference framework; hence, the results reflect relative performance assessment rather than independent validation against in situ observations. Importantly, these experiments provide a first, controlled synthetic assessment of the feasibility of retrieving precipitation from TWS. The added value of our study is not to validate the SM2RAIN algorithm (which has been done extensively using multiple observational precipitation

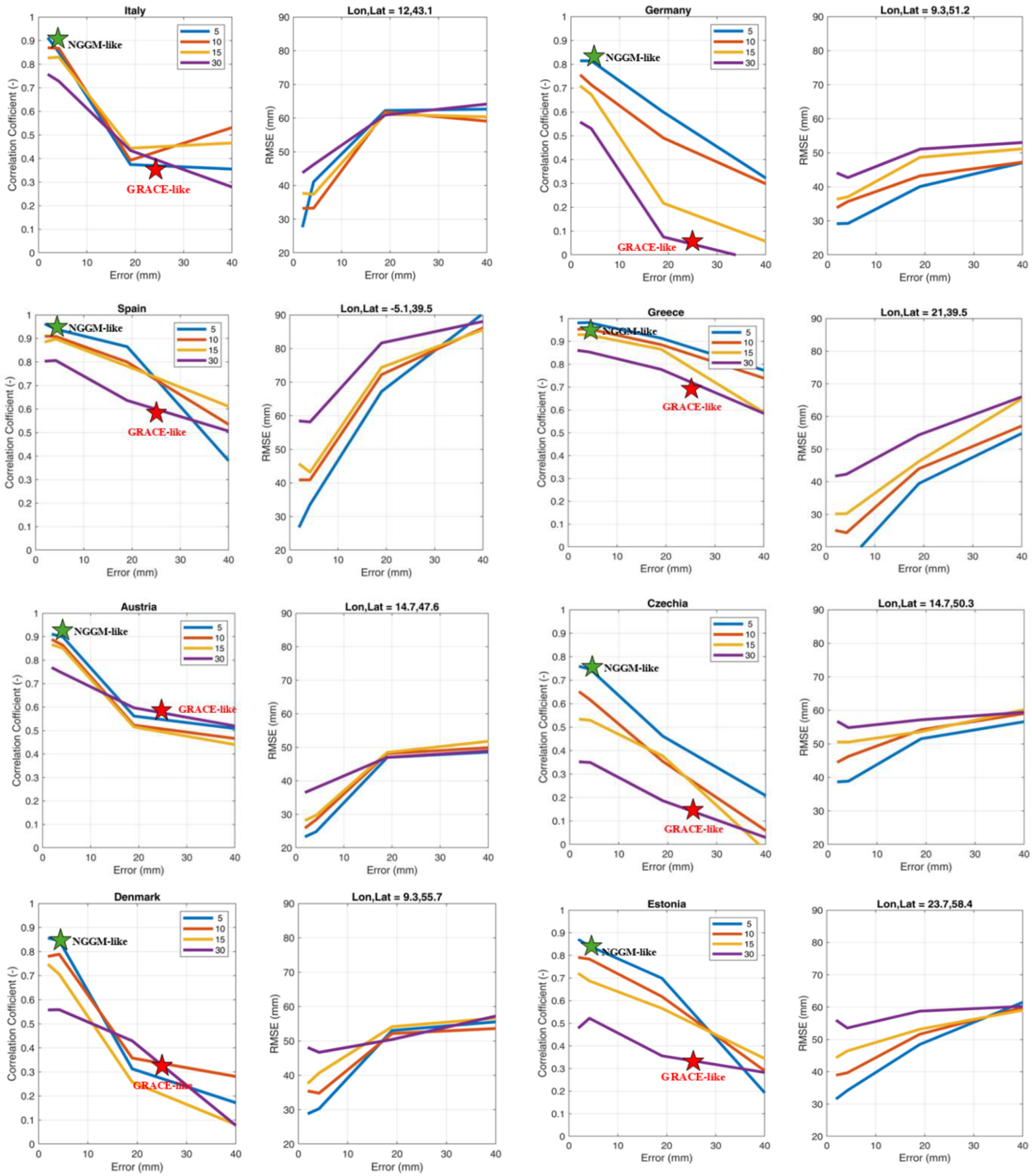

Figure 5. Evaluation of the NGGM/MAGIC synthetic experiments at four temporal resolutions (5, 10, 15 and 30 d) with error levels ranging from 0 to 40 mm for estimating 30 d precipitation over random locations in Italy, Germany, Spain, Greece, Austria, Czechia, Denmark, Estonia for the period 2003–2012 in terms of correlation coefficient (R) and root mean square error (RMSE).

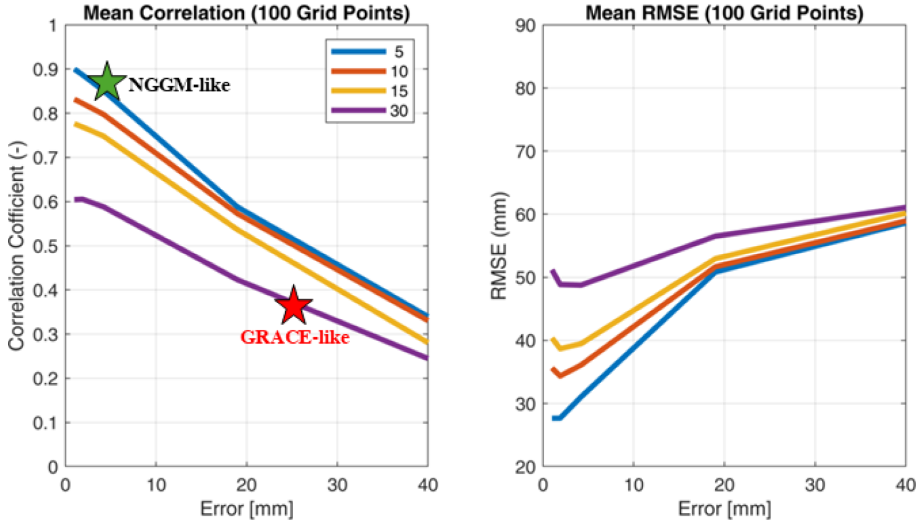

Figure 6. Evaluation of the NGGM/MAGIC synthetic experiments at four temporal resolutions (5, 10, 15 and 30 d) with error levels ranging from 0 to 40 mm for estimating 30 d precipitation over 100 random points in Europe for the period 2003–2012 in terms of correlation coefficient (R) and root mean square error (RMSE).

datasets such as e.g., GPCC or E-OBS) but rather to assess its sensitivity to the sampling and accuracy expected from future gravity missions.

Since the complete ERA5L time series was used for both calibration and evaluation, the reported performance metrics (e.g., R, and RMSE) reflect in sample agreement within this controlled synthetic framework. These statistics should therefore not be interpreted as measures of independent predictive skill or model generalization. Rather, they are intended to quantify the relative sensitivity of the proposed approach to different temporal sampling frequencies and measurement-error scenarios under internally consistent conditions. Evaluating the framework against observational datasets using independent temporal and spatial cross validation remains an important direction for future research.

Measurement uncertainties are modeled as additive Gaussian noise to perform a first-order sensitivity analysis, with standard deviations derived from NGGM mission error specifications (Daras et al., 2023). These experiments are intended as first-order sensitivity analyses rather than mission-realistic error simulations. This notice acknowledges that real satellite gravimetry errors may be temporally and spatially correlated and may be geographically dependent, with factors including latitude, topography and orbit geometry. While the Gaussian noise assumption provides a simplistic error model, it is sufficient in providing a first-order sensitivity of how measurement uncertainty may affect precipitation retrieval. In this work, our aim is not to replicate the full complexity of mission error characteristics but rather to study the relative effect of spatiotemporal sampling and measurement uncertainty on precipitation estimation from TWS. However, this comparison would demand mission specific error covariance models, the derivation of which is beyond the scope of this study.

Additionally, the SM2RAIN framework does not show evapotranspiration as independent fluxes, which could affect how storage dynamics change over weeks. Neglecting evaporation can lead to biased precipitation estimates, particularly in warm, vegetated, and semiarid regions where evapotranspiration is significant. Further, TWS changes generally include contributions from groundwater, but the current framework does not make a clear distinction between groundwater storage and other hydrological components, limiting process-level interpretation.

Future work will address these issues by integrating flux-consistent water balance constraints (incorporating evapotranspiration and runoff), parameter estimation using machine learning, and improving the attribution of groundwater dynamics with future gravity missions. This preserves the physical SM2RAIN structure while reducing calibration dependency, improving spatial coherence, and enhancing robustness in data scarce and nonstationary environments. The framework will also be extended to larger domains using more realistic, spatially, and temporally correlated error structures derived from advanced mission simulations.

5 Conclusions

Future gravity missions could play a key role in addressing the challenges of interpreting different hydrological components, which remain difficult to capture with the current Gravity Recovery and Climate Experiment (GRACE) and its Follow-On (GRACE-FO) mission. In this study, we aim to address these challenges by providing a comprehensive assessment of upcoming Next-Generation Gravity Mission (NGGM) and Mass-change And Geosciences International Constellation (MAGIC) for precipitation estimation on a European scale. The research applied the SM2RAIN algorithm by inverting the soil water balance equation and for the first time, by using the Terrestrial Water Storage (derived from the four soil layers of ERA5L soil moisture as a proxy) for precipitation estimation. The following conclusions can be drawn from the present study.

SM2RAIN algorithm exhibited satisfactory performance to simulate precipitation using TWS data across most of Europe achieving a high correlation of 0.86 and a low RMSE of 13.31 mm for 15 d temporal resolution assessments (see Fig. 2). The performance of SM2RAIN varied regionally, with stronger correlations found in the Mediterranean region and in the southern Europe, while lower values were observed in the Alpine area in northern Italy, the northeastern countries and the coastal regions of Norway, where snow processes, complex topography, and deeper storage dynamics likely affect the relationship between TWS variations and precipitation.

The verification of SM2RAIN leads to designing NGGM/-MAGIC configurations, which clearly show how temporal resolution and error levels have an impact on the performance for precipitation estimation. The results indicate high correlation values (between 0.61 and 0.88) and low RMSE values (between 28 and 50 mm) for short temporal aggregation (5 d) and for retrieval errors lower than 20 mm (see Fig. 6). For longer temporal aggregation (e.g. 30 d, which is representative of GRACE-FO-like conditions), correlation values drop significantly ranging from 0.40 and 0.60 for retrieval errors lower than 20 mm. For error levels lower than 10 mm, and temporal aggregation lower than 15 d, the performance deterioration is limited; whereas for error levels larger than 20 mm and 30 d aggregation the performance drop is highly significant. These results highlight the importance of minimizing the retrieval errors in future gravity missions. The use of TWS instead of surface soil moisture observations (e.g. Brocca et al., 2014, Brocca et al., 2019) provides a more integrated representation of the terrestrial water cycle by combining contributions from soil moisture, groundwater, and surface water storage. While soil moisture measurements typically represent only the near-surface layer, gravity-based TWS observations capture deeper storage components and large-scale water mass redistribution. Moreover, precipitation estimation from TWS may be complementary to soil moisture based approaches for large spatial scales and over regions where surface soil moisture observations are sparse and/or strongly impacted by local processes (e.g. densely vegetated areas and frozen\snow-covered soils).

The findings of this study provide critical implications for the development and execution of future gravity missions such as NGGM and MAGIC. The results suggest that TWS variations contain sufficient hydrological information to retrieve precipitation signals and that the TWS-based inversion approach could serve as a useful complement to soil moisture based precipitation estimation. The results strongly advocate the value of high temporal resolution (5 d or better) encompassing low error levels (less than 10 mm) to ensure precise precipitation estimation at both regional and global scale. The results also open new opportunities to estimate precipitation for monitoring hydrological processes from TWS measurements especially in data scarce regions. By improving spatio-temporal resolution and accuracy of gravity missions we can enhance our capacity to estimate precipitation globally, with also complementary performance with respect to the use of soil moisture measurements. This important aspect will be further explored in future studies.

Data availability. All raw and processed data can be made available by the corresponding authors upon request.

Supplement. The supplement related to this article is available online at https://doi.org/10.5194/hess-30-4969-2026-supplement.

Author contributions. LB and MUL designed the study; MUL conducted the analyses; MUL and LB interpreted the analyses; MUL wrote the manuscript draft; MUL, SC, GJ, FL and LB reviewed and edited the manuscript.

Competing interests. The contact author has declared that none of the authors has any competing interests.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. The authors bear the ultimate responsibility for providing appropriate place names. Views expressed in the text are those of the authors and do not necessarily reflect the views of the publisher.

Acknowledgements. The authors acknowledge funding from the Italian Space Agency (ASI) "NGGM/MAGIC, a breakthrough in understanding Earth's dynamics" project (grant agreement no. 2023-22.HH.0.)

Financial support. This research has been supported by the Agenzia Spaziale Italiana (grant no. 2023-22.HH.0).

Review statement. This paper was edited by Marie-Claire ten Veldhuis and reviewed by Dimos Touloumidis and Vagner Ferreira.

References

Ahrens, B.: Distance in spatial interpolation of daily rain gauge data, Hydrol. Earth Syst. Sci., 10, 197–208, https://doi.org/10.5194/hess-10-197-2006, 2006.

Ansari, R., Liaqat, M. U., and Grossi, G.: Evaluation of gridded datasets for terrestrial water budget assessment in the Upper Jhelum River Basin-South Asia, J. Hydrol.- Reg. Stud., 613, 128294, https://doi.org/10.1016/j.ejrh.2024.102114, 2022.

Beck, H. E., McVicar, T. R., Vergopolan, N., Berg, A., Lutsko, N. J., Dufour, A., Zeng, Z., Jiang, X., Van Dijk, A. I., and Miralles, D. G.: High-resolution (1 km) Köppen-Geiger maps for 1901–2099 based on constrained CMIP6 projections, Sci. Data, 10, 724, https://doi.org/10.1038/s41597-023-02549-6, 2023.

Brocca, L., Moramarco, T., Melone, F., and Wagner, W.: A new method for rainfall estimation through soil moisture observations, Geophys. Res. Lett., 40, 853–858, https://doi.org/10.1002/grl.50173, 2013.

Brocca, L., Ciabatta, L., Massari, C., Moramarco, T., Hahn, S., Hasenauer, S., Kidd, R., Dorigo, W., Wagner, W., and Levizzani, V.: Soil as a natural rain gauge: Estimating global rainfall from satellite soil moisture data, J. Geophys. Res.-Atmos., 119, 5128–5141, https://doi.org/10.1002/2014JD021489, 2014.

Brocca, L., Massari, C., Ciabatta, L., Moramarco, T., Penna, D., Zuecco, G., Pianezzola, L., Borga, M., Matgen, P., and Martínez-Fernández, J.: Rainfall estimation from in situ soil moisture observations at several sites in Europe: an evaluation of the SM2RAIN algorithm, J. Hydrol. Hydromech., 201–209, https://doi.org/10.1515/johh-2015-0016, 2015.

Brocca, L., Pellarin, T., Crow, W. T., Ciabatta, L., Massari, C., Ryu, D., Su, C., Rüdiger, C., and Kerr, Y.: Rainfall estimation by inverting SMOS soil moisture estimates: A comparison of different methods over Australia, J. Geophys. Res.-Atmos., 121, 12–62, https://doi.org/10.1002/2016JD025382, 2016.

Brocca, L., Filippucci, P., Hahn, S., Ciabatta, L., Massari, C., Camici, S., Schüller, L., Bojkov, B., and Wagner, W.: SM2RAIN–ASCAT (2007–2018): global daily satellite rainfall data from ASCAT soil moisture observations, Earth Syst. Sci. Data, 11, 1583–1601, https://doi.org/10.5194/essd-11-1583-2019, 2019.

Chua, Z.-W., Kuleshov, Y., Watkins, A. B., Choy, S., and Sun, C.: A comparison of various correction and blending techniques for creating an improved satellite-gauge rainfall dataset over Australia, Remote Sens., 14, 261, https://doi.org/10.3390/rs14020261, 2022.

Croteau, M. J., Nerem, R. S., Loomis, B. D., and Sabaka, T. J.: Development of a daily GRACE mascon solution for terrestrial water storage, J. Geophys. Res.-Sol. Earth, 125, https://doi.org/10.1029/2019JB018468, 2020.

Daras, I. (Ed.): Next Generation Gravity Mission (NGGM): Mission Requirements Document, Issue 1.0, Earth and Mission Science Division, European Space Agency, https://doi.org/10.5270/ESA.NGGM-MRD.2023-09-v1.0, 2023.

Daras, I., March, G., Pail, R., Hughes, C. W., Braitenberg, C., Güntner, A., Eicker, A., Wouters, B., Heller-Kaikov, B., Pivetta, T., and Pastorutti, A.: Mass-change And Geosciences International Constellation (MAGIC) expected impact on science and applications, Geophys. J. Int., https://doi.org/10.1093/gji/ggad472, 2024.

Daras, I., March, G., Pail, R., Hughes, C. W., Braitenberg, C., Güntner, A., Eicker, A., Wouters, B., Heller-Kaikov, B., and Pivetta, T.: Mass-change And Geosciences International Constellation (MAGIC) expected impact on science and applications, Geophys. J. Int., 236, 1288–1308, https://doi.org/10.1093/gji/ggad472, 2024.

Dorigo, W., Dietrich, S., Aires, F., Brocca, L., Carter, S., Cretaux, J.-F., Dunkerley, D., Enomoto, H., Forsberg, R., and Güntner, A.: Closing the water cycle from observations across scales: Where do we stand?, Bull. Am. Meteorol. Soc., 102, E1897–E1935, 2021.

Drinkwater, M. R., Floberghagen, R., Haagmans, R., Muzi, D., and Popescu, A.: VII: Closing session: GOCE: ESA's first earth explorer core mission, Space Sci. Rev., 108, 419–432, 2003.

Famiglietti, J. S. and Wood, E. F.: Multiscale modeling of spatially variable water and energy balance processes, Water Resour. Res., 30, 3061–3078, 1994.

Filippucci, P., Brocca, L., Quast, R., Ciabatta, L., Saltalippi, C., Wagner, W., and Tarpanelli, A.: High-resolution (1 km) satellite rainfall estimation from SM2RAIN applied to Sentinel-1: Po River basin as a case study, Hydrol. Earth Syst. Sci., 26, 2481–2497, https://doi.org/10.5194/hess-26-2481-2022, 2022.

Haagmans, R. and Tsaoussi, L.: Next Generation Gravity Mission as a Mass-change And Geosciences International Constellation (MAGIC) Mission Requirements Document, Earth and Mission Science Division, European Space Agency, https://doi.org/10.5270/esa.nasa.magic-mrd.2020, 2020.

Hengl, T., Miller, M. A. E., Križan, J., Shepherd, K. D., Sila, A., Kilibarda, M., Antonijević, O., Glušica, L., Dobermann, A., and Haefele, S. M.: African soil properties and nutrients mapped at 30 m spatial resolution using two-scale ensemble machine learning, Sci. Rep., 11, 6130, https://doi.org/10.1038/s41598-021-85639-y, 2021.

Herrnegger, M., Senoner, T., and Nachtnebel, H.-P.: Adjustment of spatio-temporal precipitation patterns in a high Alpine environment, J. Hydrol., 556, 913–921, https://doi.org/10.1016/j.jhydrol.2016.04.068, 2018.

Islam, Z.: Soil loss assessment by RUSLE in the cloud-based platform (GEE) in Nigeria, Model Earth Syst. Environ., 8, 4579–4591, https://doi.org/10.1007/s40808-022-01467-7, 2022.

Kidd, C., Becker, A., Huffman, G. J., Muller, C. L., Joe, P., Skofronick-Jackson, G., and Kirschbaum, D. B.: So, how much of the Earth's surface is covered by rain gauges?, Bull. Am. Meteorol. Soc., 98, 69–78, https://doi.org/10.1175/BAMS-D-14-00283.1, 2017.

Knoben, W. J. M., Freer, J. E., and Woods, R. A.: Technical note: Inherent benchmark or not? Comparing Nash–Sutcliffe and Kling–Gupta efficiency scores, Hydrol. Earth Syst. Sci., 23, 4323–4331, https://doi.org/10.5194/hess-23-4323-2019, 2019.

Lai, Y., Tian, J., Kang, W., Gao, C., Hong, W., and He, C.: Rainfall estimation from surface soil moisture using SM2RAIN in cold mountainous areas, J. Hydrol., 606, 127430, https://doi.org/10.1016/j.jhydrol.2022.127430, 2022.

Landerer, F. W., Flechtner, F. M., Save, H., Webb, F. H., Bandikova, T., Bertiger, W. I., Bettadpur, S. V, Byun, S. H., Dahle, C., and Dobslaw, H.: Extending the global mass change data record: GRACE Follow-On instrument and science data performance, Geophys. Res. Lett., 47, e2020GL088306, https://doi.org/10.1029/2020GL088306, 2020.

Levizzani, V., Kidd, C., Kirschbaum, D. B., Kummerow, C. D., Nakamura, K., and Turk, F. J.: Satellite precipitation measurement, Springer, https://doi.org/10.1007/978-3-030-35798-6, 2020.

Li, Y., Wang, Q. J., He, H., Wu, Z., and Lu, G.: A method to extend temporal coverage of high quality precipitation datasets by calibrating reanalysis estimates, J. Hydrol., 581, 124355, https://doi.org/10.1016/j.jhydrol.2019.124355, 2020.

Liaqat, M. U., Grossi, G., Hasson, S. ul, and Ranzi, R.: Characterization of interannual and seasonal variability of hydro-climatic trends in the Upper Indus Basin, Theor. Appl. Climatol., 147, 1163–1184, https://doi.org/10.1007/s00704-021-03850-3, 2022.

McMillan, H., Jackson, B., Clark, M., Kavetski, D., and Woods, R.: Rainfall uncertainty in hydrological modelling: An evaluation of multiplicative error models, J. Hydrol., 400, 83–94, https://doi.org/10.1016/j.jhydrol.2011.01.026, 2011.

Meles, M. B., Goodrich, D. C., Unkrich, C. L., Gupta, H. V, Burns, I. S., Hirpa, F. A., Razavi, S., and Guertin, D. P.: Rainfall distributional properties control hydrologic model parameter importance., J. Hydrol.-Reg. Stud., 51, 101662, https://doi.org/10.1016/j.ejrh.2024.101662, 2024.

Moges, D. M., Kmoch, A., and Uuemaa, E.: Application of satellite and reanalysis precipitation products for hydrological modeling in the data-scarce Porijõgi catchment, Estonia, J. Hydrol.-Reg. Stud., 41, 101070, https://doi.org/10.1016/j.ejrh.2022.101070, 2022.

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D. G., Piles, M., Rodríguez-Fernández, N. J., Zsoter, E., Buontempo, C., and Thépaut, J.-N.: ERA5-Land: a state-of-the-art global reanalysis dataset for land applications, Earth Syst. Sci. Data, 13, 4349–4383, https://doi.org/10.5194/essd-13-4349-2021, 2021.

Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models part I – A discussion of principles, J. Hydrol., 10, 282–290, https://doi.org/10.1016/0022-1694(70)90255-6, 1970.

Paredes-Trejo, F., Barbosa, H., and dos Santos, C. A. C.: Evaluation of the performance of SM2RAIN-derived rainfall products over Brazil, Remote Sens., 11, 1113, https://doi.org/10.3390/rs11091113, 2019.

Satgé, F., Hussain, Y., Molina-Carpio, J., Pillco, R., Laugner, C., Akhter, G., and Bonnet, M.: Reliability of SM2RAIN precipitation datasets in comparison to gauge observations and hydrological modelling over arid regions, Int. J. Climatol., 41, https://doi.org/10.1002/joc.6704, 2020.

Save, H., Bettadpur, S., and Tapley, B. D.: High-resolution CSR GRACE RL05 mascons, J. Geophys. Res.-Sol. Earth, 121, 7547–7569, https://doi.org/10.1002/2016JB013007, 2016.

Scanlon, B. R., Zhang, Z., Save, H., Sun, A. Y., Müller Schmied, H., Van Beek, L. P., Wiese, D. N., Wada, Y., Long, D., Reedy, R. C., and Longuevergne, L.: Global models underestimate large decadal declining and rising water storage trends relative to GRACE satellite data, Proc. Natl. Acad. Sci. USA, 115, E1080–E1089, https://doi.org/10.1073/pnas.1704665115, 2018.

Song, X., Wu, P., Gao, X., Yao, J., Zou, Y., Zhao, X., Siddique, K. H. M., and Hu, W.: Rainwater collection and infiltration (RWCI) systems promote deep soil water and organic carbon restoration in water-limited sloping orchards, Agric. Water Manag., 242, 106400, https://doi.org/10.1016/j.agwat.2020.106400, 2020.

Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S., and Hsu, K.: A review of global precipitation data sets: Data sources, estimation, and intercomparisons, Rev. Geophys., 56, 79–107, https://doi.org/10.1002/2017RG000574, 2018.

Tapley, B. D., Bettadpur, S., Watkins, M., and Reigber, C.: The gravity recovery and climate experiment: Mission overview and early results, Geophys. Res. Lett., 31, https://doi.org/10.1029/2004GL019920, 2004.

Tarpanelli, A., Massari, C., Ciabatta, L., Filippucci, P., Amarnath, G., and Brocca, L.: Exploiting a constellation of satellite soil moisture sensors for accurate rainfall estimation, Adv. Water Resour., 108, 249–255, https://doi.org/10.1016/j.advwatres.2017.08.010, 2017.

Trautmann, T., Koirala, S., Carvalhais, N., Eicker, A., Fink, M., Niemann, C., and Jung, M.: Understanding terrestrial water storage variations in northern latitudes across scales, Hydrol. Earth Syst. Sci., 22, 4061–4082, https://doi.org/10.5194/hess-22-4061-2018, 2018.

Vrochidou, A.-E., Tsanis, I. K., Grillakis, M. G., and Koutroulis, A. G.: The impact of climate change on hydrometeorological droughts at a basin scale, J. Hydrol., 476, 290–301, https://doi.org/10.1016/j.jhydrol.2012.10.046, 2013.

Williams, G. P.: Friends Don't Let Friends use Nash-Sutcliffe Efficiency (NSE) or KGE for Hydrologic Model Accuracy Evaluation: A rant with data and suggestions for better practice, Environ. Model. Softw., 106665, https://doi.org/10.1016/j.envsoft.2025.106665, 2025.

Yu, Y., Hua, T., Chen, L., Zhang, Z., and Pereira, P.: Divergent changes in vegetation greenness, productivity, and rainfall use efficiency are characteristic of ecological restoration towards high-quality development in the Yellow River Basin, China, Engineering, 34, 109–119, https://doi.org/10.1016/j.eng.2023.07.012, 2024.

Yuan, S., Wang, Y., Quiring, S. M., Ford, T. W., and Houston, A. L.: A sensitivity study on the response of convection initiation to in situ soil moisture in the central United States, Clim. Dyn., 54, 2013–2028, https://doi.org/10.1007/s00382-019-05098-0, 2020.

Zhong, Y., Tian, B., Kim, H., Yuan, X., Liu, X., Zhu, E., Wu, Y., and Wang, L.: Over 60 % precipitation transformed into terrestrial water storage in global river basins from 2002 to 2021, Commun. Earth Environ., 6, 53, https://doi.org/10.1038/s43247-024-01967-7, 2025.