



Rethinking the effect of groundwater storage resolution in river discharge estimates: how baseflow contribution governs the benefits of high-resolution data

Francesco Leopardi^{1,2}, Ehsan Forootan^{3,4}, Jacopo Dari^{1,2}, Carla Saltalippi¹, Luca Brocca², Stefania Camici²

¹Department of Civil and Environmental Engineering, University of Perugia, Perugia, Italy

²Institute of Geosciences, National Research Council, Via Madonna Alta, 126, 06128 Perugia, Italy

³Geodesy Group, Department of Sustainability and Planning, Aalborg University, Rendsburggade 14, Aalborg, 9000, Denmark

⁴School of Geographical Sciences, University of Bristol, University Rd, Bristol BS8 1SS, United Kingdom

Correspondence to: Francesco Leopardi (francesco.leopardi@dottorandi.unipg.it)

Abstract. Groundwater storage (GWS) sustains river discharge through its contribution to baseflow, particularly during low-flow periods that are becoming increasingly critical under climate change, yet it remains one of the most difficult components of the hydrological cycle to monitor. The Gravity Recovery and Climate Experiment (GRACE) and its Follow-On mission (GRACE-FO) provide large-scale observations of terrestrial water storage anomalies (TWSA), which can be used to constrain GWS estimates and improve hydrological modelling of river discharge. However, their coarse effective spatial resolution (~300 km) limits their applicability to small and medium-sized river basins. Future gravimetry missions are expected to improve the spatial resolution of gravity-derived water-storage observations, but whether this improvement translates into more accurate river discharge estimates, and under which hydrological conditions, remains unclear. Here, we investigate this question using the semi-distributed hydrological model STREAM forced with GWS anomalies (GWSA) datasets at 0.1°, 0.5°, and 3° spatial resolution. The 0.1° dataset is obtained through constrained Bayesian integration of GRACE/GRACE-FO observations into the World-Wide Water Resources Assessment (W3RA) model, while the 0.5° and 3° datasets are generated by spatially degrading the 0.1° product. The experiments encompass three European pilot basins (Ebro, Danube, and Elbe) and 24 gauging stations representing a broad spectrum of drainage areas and hydrological conditions. Our results show that drainage area alone is not a reliable indicator of the added value of higher-resolution GWSA estimates: improvements in river discharge estimates are generally modest and inconsistent across basin sizes. In contrast, a more pronounced sensitivity is observed across flow regimes, particularly under low-flow (LF) and very-low-flow (VLF) conditions, for which the median reduction in RRMSE across all gauging stations is approximately 10% and 14%, respectively, when moving from the 3° to the 0.1° GWSA product. Moreover, the benefits of higher-resolution GWSA extend to higher-flow conditions as the contribution of baseflow to total river discharge increases. These results identify baseflow contribution as a key predictor of the hydrological value of higher-resolution GWSA estimates and provide a basis for assessing where improved spatial resolution of future gravimetry missions may translate into meaningful gains in river discharge estimation.



1 Introduction

Groundwater storage (GWS) constitutes the largest reservoir of liquid freshwater on Earth and is a fundamental component of the global hydrological cycle (Cuthbert et al., 2024; Cao et al., 2026). It regulates the temporal variability of streamflow through its contribution to river baseflow, sustaining rivers during prolonged dry periods and buffering the impacts of seasonal and multi-year droughts (Gleeson et al., 2012). Consequently, GWS exerts a primary control on water availability and river ecosystem functioning, particularly during low- and very-low-flow conditions when streamflow largely depends on subsurface water reserves (Hellwig et al., 2021; Petersen-Perlman et al., 2022). Increasing groundwater abstraction and climate-driven changes in recharge are placing this resource under growing pressure. Irrigation alone accounts for more than 40% of global groundwater withdrawals (Siebert et al., 2010), while the combined effects of increasing water demand and changing climatic conditions are accelerating depletion in many regions worldwide (Jasechko et al., 2024). The resulting decline in GWS reduces baseflow contributions to rivers, exacerbating drought impacts and altering streamflow regimes. Reliable monitoring of GWS is therefore essential to understand and quantify these changes across contrasting hydrological settings (Al-Aizari et al., 2026). Yet groundwater remains one of the least observed components of the terrestrial water cycle. In situ monitoring networks are spatially sparse, unevenly distributed and unavailable in many regions of the world, limiting the possibility of monitoring groundwater dynamics consistently over large spatial domains. Recently, new lines of research have been proposed based on estimates of aquifer recharge obtained through satellite measurements of surface soil moisture (Sadeghi et al., 2019; Dari et al., 2025), but these approaches are not yet sufficiently mature to allow for the development of long-term, consistent databases based on them. A breakthrough in large-scale groundwater monitoring was achieved with the launch of the Gravity Recovery and Climate Experiment (GRACE, 2002-2017) mission and its successor GRACE Follow-On (GRACE-FO, 2018-now). By measuring temporal variations in the Earth's gravity field, these satellite missions provide estimates of terrestrial water storage anomalies (TWSA), from which GWS anomalies (GWSA) can be derived after accounting for the contributions of soil moisture, snow and surface water (Rodell et al., 2009; Mehrnegar et al., 2021; Springer et al., 2026). Over the last two decades, GRACE(-FO)-derived groundwater products have substantially improved our understanding of groundwater depletion, drought evolution, climate variability and long-term water storage changes at regional to global scales. Beyond monitoring, GRACE(-FO)-derived storage information has increasingly been incorporated into hydrological modelling frameworks for river discharge estimation. Previous studies have used GRACE observations to constrain parameter calibration and reduce model equifinality (Rakovec et al., 2016; Dembélé et al., 2020), to update model states through data assimilation techniques (Schumacher et al., 2018; Forootan et al., 2024; Retegui-Schiettekatte et al., 2025), and as a proxy of GWS to represent deep subsurface water availability controlling river baseflow (Mohanasundaram et al., 2021; Camici et al., 2022; Leopardi et al., 2026). These applications have demonstrated that satellite-derived groundwater information can improve streamflow simulations, particularly where groundwater exerts a strong control on river discharge. Nevertheless, the effective exploitation of GRACE(-FO)-derived groundwater products in hydrological applications remains constrained by the intrinsic characteristics of the observations. The monthly temporal resolution limits their ability to represent rapid hydrological



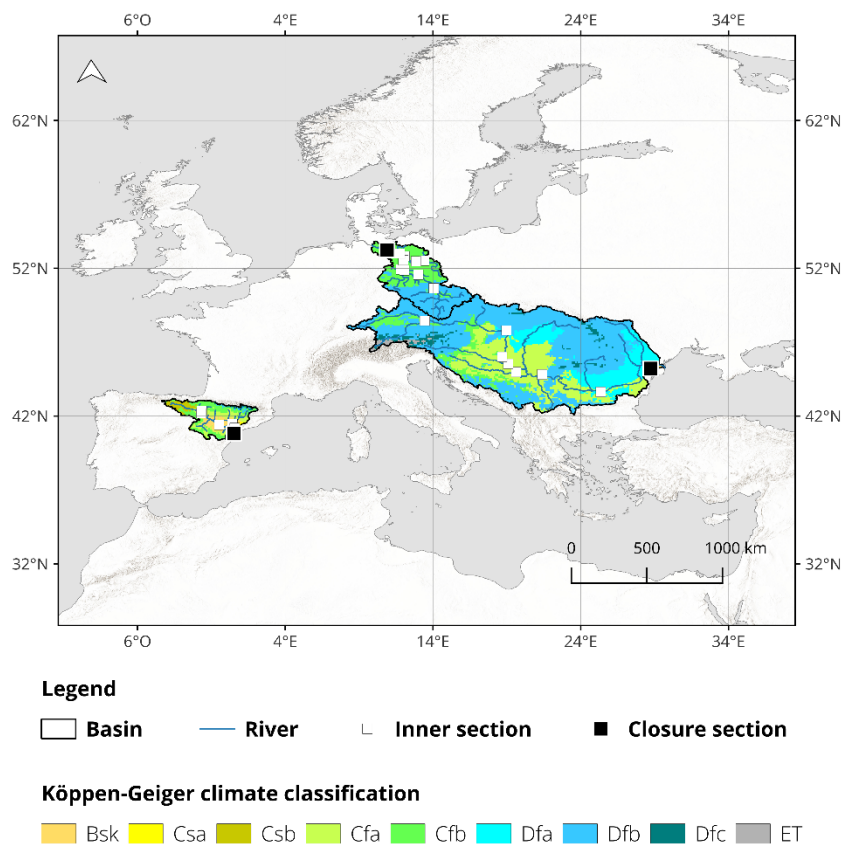
dynamics, while the effective spatial resolution of approximately 300 km and the associated spatial leakage reduce the reliability of storage estimates over medium and small river basins (Longuevergne et al., 2010; Landerer and Swenson, 2012). These limitations become particularly relevant for river discharge modelling, where hydrological processes often occur at much finer spatial scales than those resolved by current satellite gravimetry observations. To overcome the coarse native spatial resolution of satellite gravimetry, several studies have developed approaches to downscale GRACE(-FO)-derived TWSA (and hence GWSA) to finer spatial scales, ranging from statistical and regression-based methods that exploit auxiliary high-resolution predictors (Tourian et al., 2023; Mikaeeli Ahamadabad et al., 2025) to machine-learning techniques (Agarwal et al., 2023) and to dynamic downscaling through the assimilation of GRACE observations into land surface and hydrological models (Schumacher et al.; 2025, Springer et al., 2026). Downscaled GWS products have been applied to groundwater resource assessment and drought monitoring (Zhang et al., 2019; Cui et al., 2025), and, more recently, to support streamflow and baseflow estimation (Xiong et al., 2022; Xie et al., 2022; Hu et al., 2026). Substantial improvements are expected from the next generation of satellite gravimetry missions. The ESA's Next Generation Gravity Mission (NGGM) and the ESA/NASA Mass-change And Geosciences International Constellation (MAGIC) mission are designed to provide TWS observations with higher spatial resolution, improved temporal sampling and lower latency than GRACE(-FO) (Daras et al., 2024; Zingerle et al., 2024; Pfeffer et al., 2026). In particular, future missions are expected to better characterise extreme wet events at regional (Sun et al., 2024) and global scales (Middendorf et al., 2026), and to considerably broaden the use of satellite gravimetry in operational hydrology and water resources management. However, despite the anticipated improvements in future gravity missions, it remains largely unknown to what extent increasing the spatial resolution of TWSA (and hence GWSA) observations translates into more accurate river discharge simulations, particularly for medium and small catchments, where the coarse resolution of current GRACE(-FO) products represents a major limitation. In particular, no systematic assessment has investigated whether the benefits of finer-resolution groundwater information are consistent across river basins characterised by different drainage areas, hydrological regimes and groundwater contributions to streamflow. This study addresses this knowledge gap by systematically investigating the sensitivity of daily river discharge simulations to the spatial resolution of GWSA input estimates. A multi-resolution GWSA dataset derived from GRACE/GRACE-FO observations is used to explore the hydrological implications of increasing spatial resolution from 3°, representative of current GRACE mascon solutions, to 0.5°, a spatial scale relevant to both next-generation gravimetry missions and global hydrological modelling, and further to an exploratory 0.1° resolution. The analysis is performed using STREAM (SaTellite-based Runoff Evaluation And Mapping), a conceptual semi-distributed hydrological model that exploits satellite observations to estimate daily river discharge (Leopardi et al., 2026). The experiments are conducted over three European pilot basins, the Danube, Elbe and Ebro, which encompass a wide range of climatic and hydrological conditions. The objectives of this study are threefold: (i) to quantify the sensitivity of river discharge simulations to the spatial resolution of GWSA estimates; (ii) to assess whether the benefits of finer-resolution groundwater information depend on basin characteristics, including drainage area, flow regime and baseflow contribution; and (iii) to provide guidance on the expected hydrological value of future satellite gravimetry missions. By systematically isolating the effect of GWSA spatial resolution across contrasting basin and



100 hydrological conditions, this study provides a controlled framework for assessing when and where higher-resolution satellite-
derived water storage observations can improve hydrological modelling and inform water resources assessment.

2. Study areas

Three European basins are selected as case studies for this work (see Fig. 1): the Elbe river basin in Central Europe, the Ebro
river basin, located in the Iberian Peninsula in Southwestern Europe and the Danube river basin, extending across Central and
105 Southeastern Europe. These pilot basins are chosen because they differ markedly in climatic setting, drainage area and
dominant hydrological driver and response.



110 **Figure 1: Location of the selected gauging stations within the three pilot river basins, overlaid on the Köppen–Geiger climate classification.**

From a climatic perspective, and following the Köppen–Geiger classification of Beck et al. (2023), the three basins encompass
contrasting regimes. The Ebro is the most heterogeneous one: its central part combines a cold semi-arid steppe climate (Bsk)
with a temperate, warm-summer climate without a dry season (Cfb); the lower basin near the river mouth is Mediterranean,



with dry, hot summers (Csa); and the mountainous margins are cold and humid with warm summers (Dfb). The Elbe is comparatively uniform, transitioning from the temperate Cfb regime in the north to the cold Dfb regime in the south. The Danube is dominated by a temperate, hot-summer climate without a dry season (Cfa) in its central reaches, grading into cold, humid climates (Dfb) towards its Northern and Southern edges. Together, the three basins span a wide portion of the European climatic spectrum, from the arid and Mediterranean-influenced conditions of the Ebro, through the oceanic setting of the Elbe, to the more continental, hot-summer conditions of the Danube. The basins also differ in size by almost an order of magnitude. The Ebro is the smallest one, draining 84,230 km² at its outlet gauge in Tortosa; the Elbe drains 131,950 km² at Neu Darchau; and the Danube, by far the largest, drains 807,000 km² at Ceatal Izmail. Finally, the three basins differ in their hydrological response, and in particular in the degree to which the total river discharge is sustained by groundwater rather than by a direct response to precipitation. Following Beck et al. (2013; 2015), this behaviour can be characterised through the baseflow index (BFI), which reflects how much of the total annual streamflow is fed by slow, groundwater-supported drainage. In terms of mean annual flow, the three basins describe a clear gradient. The Elbe is the most groundwater-driven: its river discharge is strongly buffered by subsurface storage and responds smoothly to hydrological forcing. The Danube occupies an intermediate position, while the Ebro is the most rainfall-driven, with a flashier response and a comparatively lower baseflow contribution to total annual discharge. However, during low-flow periods baseflow becomes the dominant source sustaining river discharge in this basin, so that GWS exerts a particularly strong control precisely when water availability is most critical. This contrast, a gradient in mean-annual baseflow contribution combined with a strong groundwater control on Ebro low flows, offers a natural interpretive framework for assessing how river discharge simulations respond to the resolution of groundwater storage observations in each basin.

3. Methodology and datasets

This section describes the hydrological model adopted in this study, together with the different data sources used as meteorological forcings and for the calibration of the model parameters.

3.1 STREAM model

The analysis is performed using the STREAM hydrological model (Camici et al., 2022; Leopardi et al., 2026), a conceptual semi-distributed model designed to estimate daily river discharge by combining satellite observations with observed hydrological variables. STREAM is particularly suited for the purpose of this study because it represents river discharge as the sum of two main components: a fast-flow component, primarily driven by precipitation and superficial soil water content, and a slow-flow component, primarily controlled by GWS. This formulation enables the contribution of GWSA to river discharge to be explicitly investigated. The reader is referred to Leopardi et al. (2026) for a complete description of the model equation and parameters, while a brief overview of the model formulation adopted in this study is reported in Appendix A. The configuration adopted in this study differs from the original implementation of Leopardi et al. (2026) in terms of the input



145 data. Specifically, the Soil Water Index (SWI) at 5–15 cm depth is used here to represent superficial satellite-based soil
 moisture in the formulation of the fast-flow component. Moreover, in the original formulation, the slow-flow component was
 driven by TWSA, after correcting the signal only for the snow-storage anomaly simulated by STREAM's internal snow module
 (Leopardi et al., 2026). No further partitioning of the TWS signal was performed; thus, the resulting anomaly was implicitly
 used as a proxy for groundwater storage dynamics. In the present study, by contrast, model-based GWSA is explicitly
 150 prescribed as the input to the slow-flow component, enabling the sensitivity of simulated river discharge to GWSA spatial
 resolution to be directly assessed. The use of TWSA as a proxy for GWS in the original STREAM formulation represents a
 simplifying assumption, as TWSA integrates multiple terrestrial water storage components and is not equivalent to GWS.
 While TWSA may provide information on integrated storage dynamics at seasonal to sub-seasonal timescales (Tourian et al.,
 2018; Scanlon et al., 2019), its direct interpretation as GWS can introduce uncertainty in the representation of groundwater-
 155 related processes. By explicitly prescribing GWSA in the present study, this assumption is avoided, allowing the role of
 groundwater storage and, in particular, its spatial resolution to be investigated directly.

3.2 Datasets

This study employs datasets from different sources as forcings of the STREAM hydrological model and as reference data for
 parameter calibration and validation of river discharge estimates. The datasets used are listed in Table 1, which reports the
 160 represented variables, the datasets name, the sources — SAT for satellite-based products, MOD for model-based products, GR
 for ground-based data — the native spatial resolutions and the temporal coverages. A detailed description of each dataset is
 provided in the following sections.

Table 1. Meteo and EO data used to estimate river discharge time series. The labels SAT, MOD, and GR denote satellite-
 165 based, modelled-based and ground-based products, respectively.

#	Dataset type	Dataset name	Data source(s)	Resolution	Temporal coverage	Spatial coverage
1	Precipitation	IMERG-F v6	SAT	daily/0.1°	1998-continuing	Global
2	Soil water index	SWI	SAT	daily/0.1°	2007-continuing	Global
3	Snow cover fraction	ESA Climate Change Snow Cover Fraction	SAT	daily/0.01°	2000-2023	Global
4	Air temperature	ERA5-Land	MOD	daily/0.1°	1950-continuing	Global



5	Groundwater storage changes	GWSA	MOD	monthly/0.1°	2003- 2017; 2019-2024	Europe
6	River discharge	Global Runoff Data Centre	GR	daily/point	1900- continuing	River basin

3.2.1 Satellite Data

Three satellite-based datasets are used in this study: precipitation, soil water index and snow cover fraction. Precipitation data are provided by the Integrated Multi-satellite Retrievals (IMERG) for Global Precipitation Measurements (GPM) Final run product (IMERG-F V07, Huffman et al., 2020). IMERG-F represents the most accurate option in the IMERG product family (Pradhan et al., 2022) owing to monthly gauge corrections incorporating Global Precipitation Climatology Center (GPCC) data. Daily precipitation estimates at $0.1^\circ \times 0.1^\circ$ resolution are available from June 1998 onwards at https://gpm1.gesdisc.eosdis.nasa.gov/data/GPM_L3/GPM_3IMERGDF.07/.

Soil moisture information is obtained from the Soil Water Index version 4 (ASCAT SWI v4) dataset, which represents soil water dynamics at different depths. The SWI is derived from Advanced SCATterometer (ASCAT) surface soil moisture observations using a two-layer water balance model proposed by Wagner et al. (1999). It provides global daily estimates of the water content in the surface soil layer at $0.1^\circ \times 0.1^\circ$ spatial resolution from 2007 onwards. The dataset is freely available at: <https://land.copernicus.eu/en/products/soil-moisture/daily-soil-water-index-global-12-5km-v4>. For this study, a SWI with a characteristic time length (T) equal to 10 days is adopted as a proxy for the water content in the upper 5–15 cm of the soil profile.

Snow cover fraction (SCF) data are provided by the ESA Snow Climate Change Initiative (Snow CCI) project, version 4. The dataset offers global daily estimates of SCF at $0.01^\circ \times 0.01^\circ$ spatial resolution for the period 2000–2023, and is freely available at: <https://catalogue.ceda.ac.uk/uuid/375ffdb8f0a445e380b4b9548655f5f9/>.

3.2.2 Modelled Data

Two modelled-based datasets are used in this study: air temperature and GWSA. Air temperature data are provided by the ERA5-Land (European ReAnalysis v5-Land) reanalysis (Muñoz-Sabater et al., 2021). ERA5-Land is produced by the European Centre for Medium-Range Weather Forecasts (ECMWF) and offers global estimates of 2 m air temperature above the land surface from 1950 to the present, at daily temporal resolution and 0.1° spatial resolution. The dataset is freely available at <https://cds.climate.copernicus.eu/>.

The GWSA data used to force the STREAM model at its native, finest spatial resolution are obtained through an extended Constrained Bayesian (ConBay) integration framework (Mehrnegar et al., 2021; 2023; 2026), which combines satellite gravimetric observations from GRACE (2003–2017) and GRACE-FO (2018–present) with outputs from the W3RA hydrological model. Prior to assimilation, W3RA total water storage fields are transformed into the spectral domain and filtered using GRACE-consistent operators; after Bayesian updating, residual components are reintroduced to enable physically



195 coherent spatial downscaling of the terrestrial water storage signal. A similar formulation has been applied to characterise
 groundwater drought severity across Europe over the period 2003–2021 (Forootan et al., 2024). The resulting GWSA product
 used in this study covers the European domain at approximately 10 km (0.1°) spatial resolution and monthly temporal
 resolution, and represents the finest spatial configuration examined in this work. Because of the observational gap between the
 GRACE and GRACE-FO missions, no GWS data are available between 2017 and 2018; the present study therefore restricts
 200 the analysis to the 2007–2017 period (Sect. 4). The dataset is freely available at
<https://doi.org/10.6084/m9.figshare.31073527.v2>.

3.2.3 In situ Data

The only source of ground-based data is represented by river discharge observations, used for calibration and validation
 purposes. Twenty-four daily river discharge records for the three study basins, collected at the gauging stations listed in Table
 205 2, are obtained for the 2003–2022 period from the Global Runoff Data Centre (GRDC) portal at <https://grdc.bafg.de/> and from
 the Sistema Automático de Información Hidrológica (SAIH) Ebro portal at <https://saihebro.com/>.

Table 2. List of gauging stations used in this study. The table presents the station name and its respective river basin, station
 coordinates, and drainage areas. Stations located at the outlets of the investigated basins are shown in bold.

#	Station name	River Basin	Latitude	Longitude	Drainage area (km ²)
1	Caparroso	Ebro	42.34	−1.65	5469
2	Seros	Ebro	41.45	0.42	12782
3	Castejon	Ebro	42.18	−1.69	25194
4	Gelsa	Ebro	41.40	−0.48	47133
5	Ascò	Ebro	41.19	0.57	82245
6	Tortosa	Ebro	40.81	0.52	84230
7	Schaerding	Danube	48.44	13.44	25520
8	Sremska Mitrovica	Danube	44.98	19.62	87966
9	Nagymaros	Danube	47.79	18.96	223400
10	Bogojevo	Danube	45.53	19.08	251593
11	Bazias	Danube	44.81	21.38	570896
12	Zimnicea	Danube	43.63	25.35	658400



13	Mohacs	Danube	46.00	18.68	251300
14	Ceatal Izmail	Danube	45.22	28.72	807000
15	Berlin-Mühlendamm Op	Elbe	52.51	13.41	9707
16	Ketzin	Elbe	52.48	12.85	16173
17	Calbe Grizehne	Elbe	51.92	11.81	23719
18	Havelberg Stadt	Elbe	52.82	12.08	23858
19	Ústí nad Labem	Elbe	50.65	14.04	48560
20	Torgau	Elbe	51.56	13.01	55211
21	Aken	Elbe	51.86	12.06	70093
22	Tangermünde	Elbe	52.55	11.98	97780
23	Wittenberge	Elbe	52.99	11.76	123532
24	Neu Darchau	Elbe	53.23	10.89	131950

210 4. Experimental design and performance metrics

The analyses carried out aim to identify the basin characteristics for which changes in GWSA spatial resolution, from coarser to finer spatial scales, lead to measurable differences in simulated river discharge. To this end, the analysis follows a three-step framework: (i) generation of GWSA datasets at different spatial resolutions, (ii) calibration of the STREAM model using the resulting configurations, and (iii) assessment of river discharge sensitivity to GWSA spatial resolution across basin characteristics and flow conditions. The GWSA datasets and calibration procedure are described below, while the sensitivity analysis is presented in separate dedicated paragraphs.

Three GWSA spatial configurations are considered:

1. 0.1° (native) spatial resolution. The GWSA data are used at their native spatial resolution, while all other input datasets are simply remapped onto this native 0.1° GWSA grid, preserving their original spatial information.
- 220 2. 0.5° spatial resolution. The GWSA data are spatially degraded to 0.5° and then regridded back onto the native 0.1° grid. All other input datasets are kept at their native 0.1° resolution.
3. 3° spatial resolution. The GWSA data are spatially degraded to 3° and then regridded back onto the native 0.1° grid, while all other input datasets are kept at their native 0.1° resolution. This configuration approximately represents the effective spatial resolution of GRACE-derived products and is therefore adopted as the benchmark configuration.

225



Notably, the expected spatial resolution of the forthcoming NGGM/MAGIC mission is approximately 100–400 km, depending on the target application and averaging period (Daras et al., 2024), placing its anticipated performance between the spatial scales represented by the 0.5° and 3° configurations considered here. The 0.5° resolution is selected to be consistent with that commonly adopted by global hydrological models (Döll et al., 2003; Bierkens et al., 2015; Heinicke et al., 2024), whereas 3° is selected to approximate the effective spatial resolution of GRACE mascon solutions (Watkins et al., 2015). These configurations therefore span spatial scales relevant to both current and next-generation gravimetry observations. Importantly, the aim of this study is not to determine the specific spatial resolution achievable by current or future gravimetry missions, but rather to identify the basin and hydrological characteristics under which increasing the spatial resolution of gravimetry-derived GWS observations yields tangible improvements in river discharge estimation.

Starting from the native 0.1° GWSA dataset, the two coarser spatial configurations (0.5° and 3°) are obtained through a spatial degradation procedure applied independently to each monthly GWSA field. The degradation consists of a two-dimensional moving-average (box-car) smoothing (Shuai et al., 2022), implemented as the convolution of the native GWSA grid with a uniform square kernel. A 5×5 pixel kernel is used for the 0.5° configuration, whereas a 30×30 pixel kernel is applied for the 3° configuration, consistently with the native spatial resolution. To account for invalid pixels, such as those outside the study domain or at the land–sea boundary, a binary mask is applied. The convolution is then computed using this mask, so that only valid data are considered. The smoothing is therefore computed considering only valid pixels, preventing artificial values from being introduced in the degraded fields. This procedure is repeated independently for each monthly GWSA field, resulting in two synthetic coarser-resolution datasets. Such degradation procedure here adopted is not intended to reproduce independent coarse-resolution gravimetry observations. Instead, it provides a controlled experimental framework in which progressively smoothed GWSA fields are used to isolate the effect of spatial resolution on river discharge simulations while keeping the remaining modelling components unchanged. Because the GWSA product is provided at monthly temporal resolution, data is linearly interpolated to the daily scale to match the temporal resolution of the other model inputs. The complete time series, for all three configurations, are then restricted to the 2007–2017 period, to avoid the data gap between the GRACE and GRACE-FO missions.

The STREAM model is calibrated at the closure section of each basin—Ceatal Izmail for the Danube, Neu Darchau for the Elbe, and Tortosa for the Ebro—against observed daily river discharge by maximising the Kling–Gupta Efficiency (KGE; Gupta et al., 2009). The resulting parameter set is then transferred to the internal gauging stations without further optimisation. This approach avoids multiple calibrations within the same basin that could compensate for or mask the effects of GWSA spatial resolution (0.1°, 0.5°, and 3°) on simulated river discharge. The procedure is repeated for each GWSA resolution and basin, resulting in nine independent calibrations. As the objective is to assess the effect of GWSA spatial resolution rather than to optimise operational river discharge predictions, no temporal split between calibration and evaluation periods is applied. Instead, the full 2007–2017 observed river discharge time series is used for calibration, maximising the information available to assess the effect of GWSA resolution under different hydrological conditions. Another reason for not adopting a temporal split is that model calibration is performed exclusively at the basin outlet, whereas model performance is subsequently



260 evaluated at internal gauging stations using the corresponding observed river discharge series, which are not used to constrain the model parameters. This provides an independent spatial evaluation of model performance within the basin, complementing the calibration at the outlet. After calibration of the STREAM model, the simulated river discharge obtained with the three GWSA configurations is analysed according to three basin characteristics: contributing drainage area, flow regime, and baseflow contribution to total river discharge, jointly used here as complementary basin descriptors (Sawicz et al., 2011).

265 These characteristics are selected because they represent complementary controls on the propagation of groundwater storage information through the hydrological system: basin area reflects the spatial representativeness of the GWSA forcing, whereas flow regime — defined here as the relative contribution of low-, intermediate-, and high-flow conditions — and baseflow contribution describe hydrological processes that are expected to be particularly sensitive to changes in groundwater storage. Model performance is quantified using the Relative Root Mean Square Error (RRMSE), computed between simulated and

270 observed daily river discharge at each selected gauging station for the three GWSA spatial configurations. Lower RRMSE values indicate a closer agreement between simulated and observed river discharge, whereas higher values correspond to larger deviations and therefore poorer model performance. The RRMSE values obtained using the GWSA configuration at 3° spatial resolution are set as the benchmark. The percentage change between the benchmark and the results obtained with the finer spatial resolution configurations (0.1° and 0.5°) is then quantified through the relative reduction in RRMSE with respect to the

275 3° baseline (positive Δ RRMSE corresponds to a reduction in RRMSE, and hence to an improvement in model performance relative to the 3° reference, whereas a negative Δ RRMSE indicates a degradation).

4.1 Assessing the role of basin area in groundwater storage resolution sensitivity

To assess the sensitivity of river discharge estimates to GWSA spatial resolution across different basin scales, the selected gauging stations are grouped into three clusters based on their contributing drainage area. The first cluster comprises stations

280 with a drainage area below 50,000 km², deliberately set below the minimum applicability threshold identified in the literature, that is the minimum catchment area above which GRACE-derived TWS signals can be resolved at an acceptable error level (Vishwakarma et al., 2018). This conservative choice ensures that all stations in this group are characterised by contributing areas smaller than the spatial resolution of GRACE-based products, where leakage errors and spatial representativeness limitations are expected to be most severe. The second cluster includes stations with drainage areas between 50,000 and

285 200,000 km², above the minimum applicability threshold. This cluster represents a transitional group between subbasins that are too small for reliable GRACE data application and those whose drainage area substantially exceeds the spatial resolution of the data, but which may still be affected by the spatial resolution of the GWSA forcing. The third cluster comprises stations with drainage areas exceeding 200,000 km², well above the spatial resolution of the GRACE data. At these scales, the contributing area encompasses a large number of GWSA pixels even at the coarsest resolution (3°), so that the extracted signal

290 represents a spatially stable average over the basin.



4.2 Assessing the role of flow regime in groundwater storage resolution sensitivity

This section describes the experimental design adopted to investigate the impact of GWSA spatial resolution on river discharge estimates as a function of the flow regime. The behaviour observed across different flow regimes represents one of the most informative characteristics of a river basin, as it synthesises the dominant hydrological processes operating within the catchment. The method employed to characterise the different flow regimes is the empirical Cumulative Distribution Function (CDF), which relates river discharge to the probability that a given flow value is equalled or not exceeded. The CDF is computed for each selected gauging station and divided into five flow classes according to the classification proposed by Pfannerstill et al. (2014): Very Low Flow (VLF, $P \leq 0.05$), representing extreme drought conditions; Low Flow (LF, $0.05 < P < 0.20$), corresponding to low-flow conditions; Middle Flow (MF, $0.20 \leq P < 0.80$), representing ordinary flow conditions; High Flow (HF, $0.80 \leq P < 0.95$); and Very High Flow (VHF, $P \geq 0.95$), representing extreme flood conditions. This classification is particularly relevant for this study, as it allows the investigation of the impact of GWSA spatial resolution across the full range of river discharge conditions. The same segmentation is applied consistently to both the observed and simulated river discharge time series obtained using the three GWSA spatial configurations (0.1° , 0.5° and 3°). For each flow regime class, the RRMSE is computed between the simulated and observed river discharge values across all stations within each pilot basin, allowing the assessment of whether model performance varies systematically as a function of GWSA spatial resolution across the different regimes or remains uniform along the discharge distribution.

4.3 Assessing the role of baseflow contribution in groundwater storage resolution sensitivity

The analysis described in this section represents a further development of the experimental design presented in Sect. 4. This analysis focuses on the role of the baseflow contribution to total river discharge as a potential control on the sensitivity of the model to GWSA spatial resolution. For each selected gauging station within the three pilot basins, the mean annual baseflow contribution to total river discharge (BF/Q) is computed as the ratio of mean annual baseflow to mean annual total river discharge (Smakhtin, 2001). This metric provides a time-integrated representation of the relative contribution of baseflow to streamflow and is representative of the long-term hydrological behaviour of the basin, while reducing the influence of interannual variability. The BF/Q ratio is derived from STREAM simulations, since the model separates river discharge into slow-flow and fast-flow components. The relationship between ΔRRMSE and the BF/Q ratio, is then analysed to investigate whether the improvement in river discharge estimates associated with finer GWSA spatial resolution varies systematically with the contribution of the slow-flow component to total river discharge.

5. Results

This section presents the results of the analysis of the impact of GWSA spatial resolution on river discharge estimates. Following the experimental design described in Sect. 4, the results are organised according to the main factors investigated: basin area (Sect. 5.1), the flow regime (Sect. 5.2) and the baseflow contribution to total river discharge across different flow

regimes (Sect. 5.3). Although the performance of the STREAM model in reproducing observed river discharge across the three study basins is not the primary focus of this study, a brief evaluation is provided in Appendix C to document the reliability of the simulations supporting the sensitivity analysis. The reader is referred to the corresponding section of the Appendix C for further details.

5.1 Response of river discharge estimation to groundwater storage resolution across basin area

To assess the effect of GWSA spatial resolution across basin sizes, RRMSE is plotted against the contributing drainage area of each selected station. Following the clustering described in Sect. 4.1, the 24 gauging stations are grouped into three area classes: 10 stations in the first (contributing area below 50,000 km²), 8 in the second (between 50,000 and 200,000 km²), and 6 in the third (above 200,000 km²). The smaller number of stations in the largest class reflects the limited availability of European gauging stations draining areas larger than 200,000 km². Figure 2 shows the RRMSE as a function of contributing drainage area for the three GWSA configurations. In the smallest basins (<50,000 km²), RRMSE is lower for the 0.1° and 0.5° configurations, with median values of 75% compared to 79% obtained with the 3° configuration. For the second cluster (50,000–200,000 km²), the median RRMSE is approximately 64% for all three configurations. Similarly, for the third cluster (> 200,000 km²), the median RRMSE is approximately 35% across all spatial resolutions.

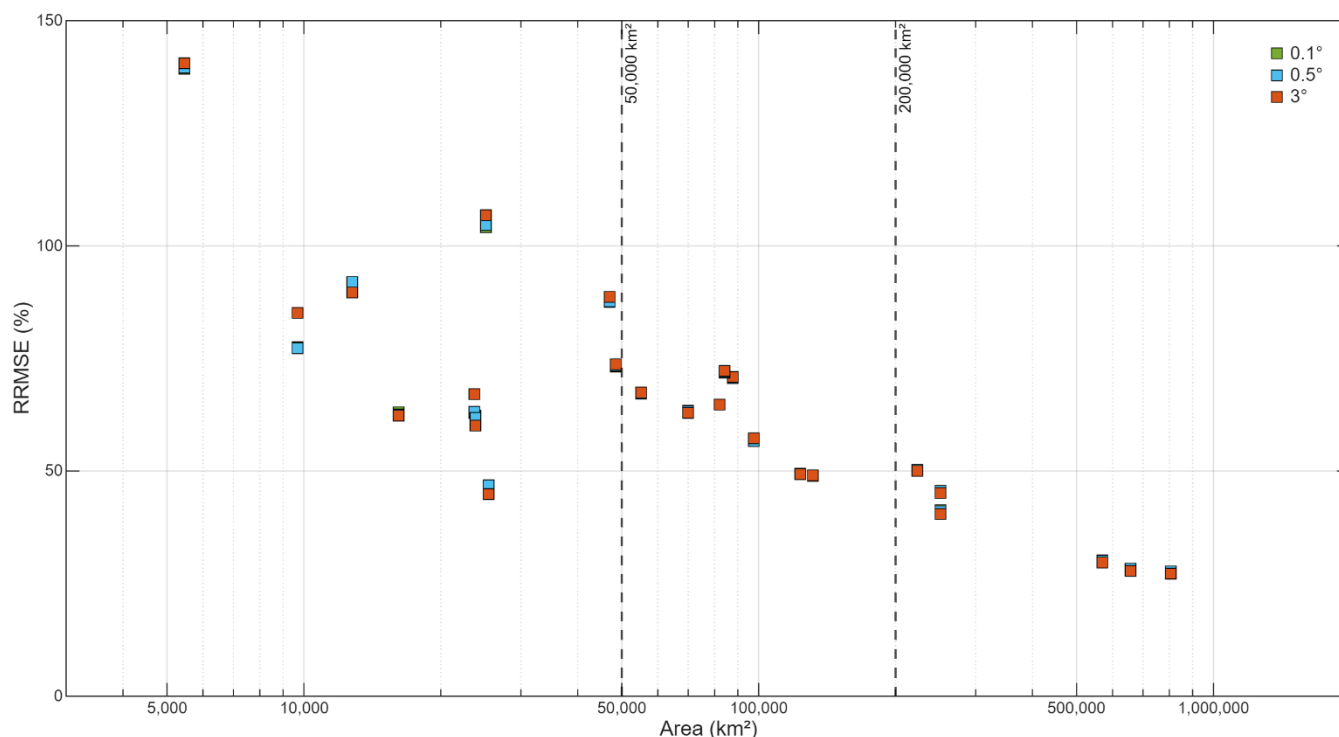


Figure 2: RRMSE (%) values for the contributing areas of the stations analysed in this study, presented for the respective spatial resolution of the GWSA data.



Figure 3 presents the Δ RRMSE as a function of the contributing drainage area. For the first cluster, the median improvement is modest in both cases, with a reduction of 0.73% when moving from 3° to 0.1° and 0.51% from 3° to 0.5°. For the second cluster, the median improvement is even smaller, with a reduction of 0.25% from 3° to 0.5° and a negligible change from 3° to 0.1°. For the third cluster, the Δ RRMSE is slightly negative, indicating a marginal increase in RRMSE of 1.73% when moving to 0.1° and 1.44% when moving to 0.5°, although the magnitude of these changes remains below 2%. Overall, the results indicate that the drainage area does not exert a systematic control on the sensitivity of river discharge estimates to the different GWSA spatial resolutions, as the observed improvements and deteriorations remain small and do not follow a consistent pattern across the three area clusters. This finding does not necessarily contradict the basin-area thresholds identified in the GRACE literature (Save et al., 2016; Vishwakarma et al., 2018), rather, it suggests that, once a GWS product is available, the sensitivity of the resulting river discharge estimates to its spatial resolution is governed less by basin geometry than by the hydrological processes operating within the catchment.

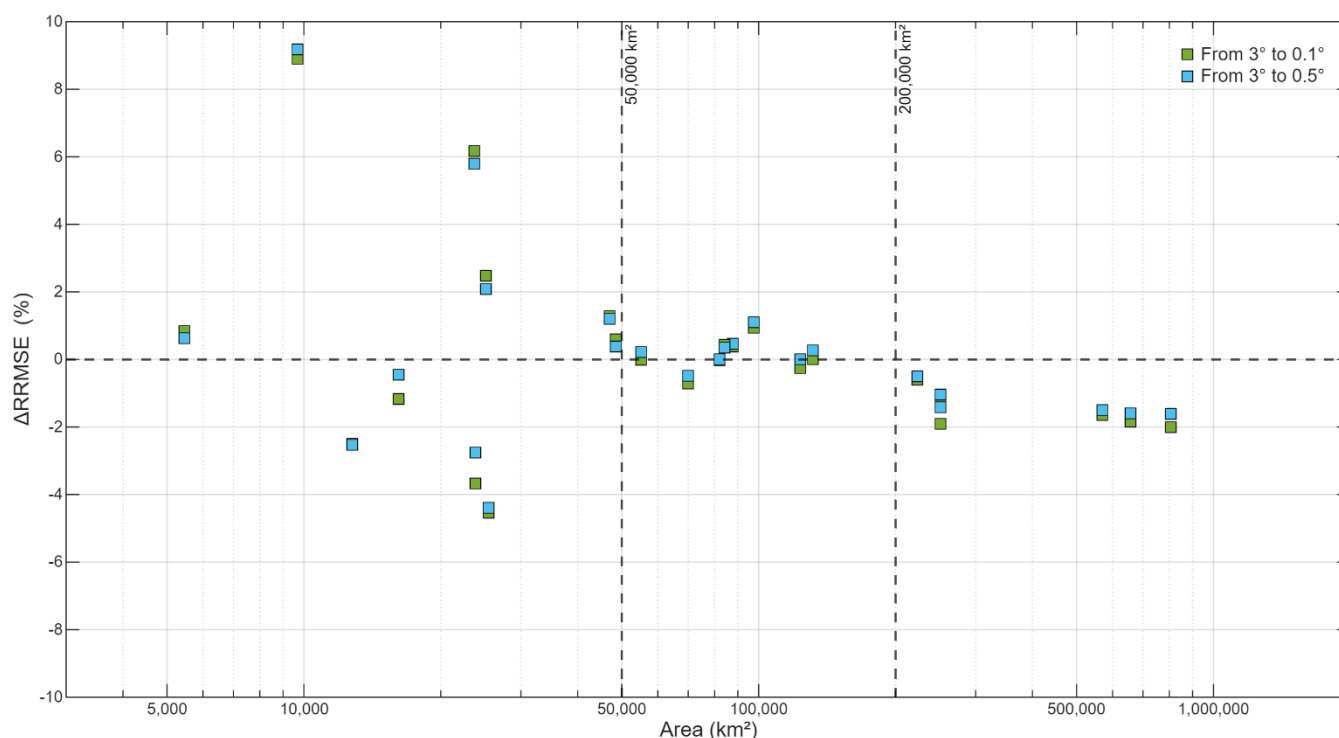


Figure 3: The figure shows the Δ RRMSE (%) values for the contributing areas of the stations analysed in this study, illustrating the changes obtained when decreasing the GWSA spatial resolution from 3° to 0.1° and 0.5°, respectively.

5.2 Response of river discharge estimation to groundwater storage resolution across flow regimes

A further analysis is conducted to investigate the impact of GWSA spatial resolution across the different flow regimes. The results are summarised in the following figures at the basin scale, aggregating the performance scores obtained at all selected stations within each pilot basin. For each basin, the RRMSE values across the five flow regimes are presented first, followed



by the corresponding Δ RRMSE. For the Ebro river basin (Fig. 4, upper panel), the highest median RRMSE values for all three configurations are recorded for the VHF regime, with values of approximately 99%, while the lowest values are observed for the LF regime, with a median RRMSE of approximately 95% for the 3° configuration and 93% for both the 0.5° and 0.1° configurations. The corresponding Δ RRMSE analysis (Fig. 5) reveals that adopting finer GWSA resolutions does not produce any reduction in the relative error for the VHF, HF and MF regimes. For the LF and VLF regimes, however, a finer spatial resolution reduces the RRMSE by approximately 2% when moving from 3° to either 0.5° or 0.1° . A similar pattern is observed for the Danube river basin (Fig. 4, middle panel), although with substantially different magnitudes. The median RRMSE values for the VHF regime reach 407% for the 3° configuration and 388% and 387% for the 0.5° and 0.1° configurations, respectively. These elevated values reflect the difficulty of the model in reproducing extreme flood events, especially in internal sections. Indeed, as shown in the Appendix C (Fig. C1, Fig. C2, Fig. C3), the model captures river discharge peaks with good accuracy at the closure stations, where the parameters were calibrated, whereas it tends to underestimate the higher river discharge values at the stations where the parameters were spatially transferred from the closure section. The lowest median RRMSE values are recorded for the LF regime (97% for the 3° configuration and approximately 70% for both finer configurations) and for the VLF regime (130% for 3° , 86% and 87% for 0.5° and 0.1° , respectively). Unlike the Ebro, the Δ RRMSE analysis for the Danube (Fig. 5) shows that a finer GWSA spatial resolution reduces the RRMSE across all flow regimes. The VHF regime remains the least sensitive, with a modest median reduction of 5% when moving from 3° to 0.1° and 0.5° . The most pronounced improvements are observed for the VLF regime, with a median reduction of approximately 22% when moving from the 3° baseline to the 0.1° configuration and 23% when moving to the 0.5° configuration. The Elbe river basin (Fig. 4, lower panel) confirms the pattern observed in the previous two basins. The highest median RRMSE values are recorded for the VHF regime (approximately 95% for the 3° configuration and 96% for both finer configurations), while the lowest values for the 0.1° and 0.5° configurations are observed for the LF regime and for the 3° configuration in the HF regime. The Δ RRMSE analysis (Fig. 5) shows negative median values for the VHF regime (approximately -0.9% for both finer configurations), confirming that the VHF regime is the least sensitive to changes in GWSA spatial resolution. The most pronounced improvement is again observed for the VLF regime, with a median RRMSE reduction of 14% when moving from 3° to 0.1° and 10% from 3° to 0.5° . Across the three pilot basins, a consistent pattern emerges: the impact of GWSA spatial resolution on river discharge estimates is systematically more pronounced for the LF and VLF regimes, while the VHF and HF regimes show minimal or no sensitivity to the GWSA spatial resolution. The magnitude of the improvement varies across basins, with the Danube exhibiting the largest Δ RRMSE values and the Ebro the smallest (see Fig. 5), suggesting that additional basin-specific characteristics beyond the flow regime may modulate the sensitivity to GWSA spatial resolution.

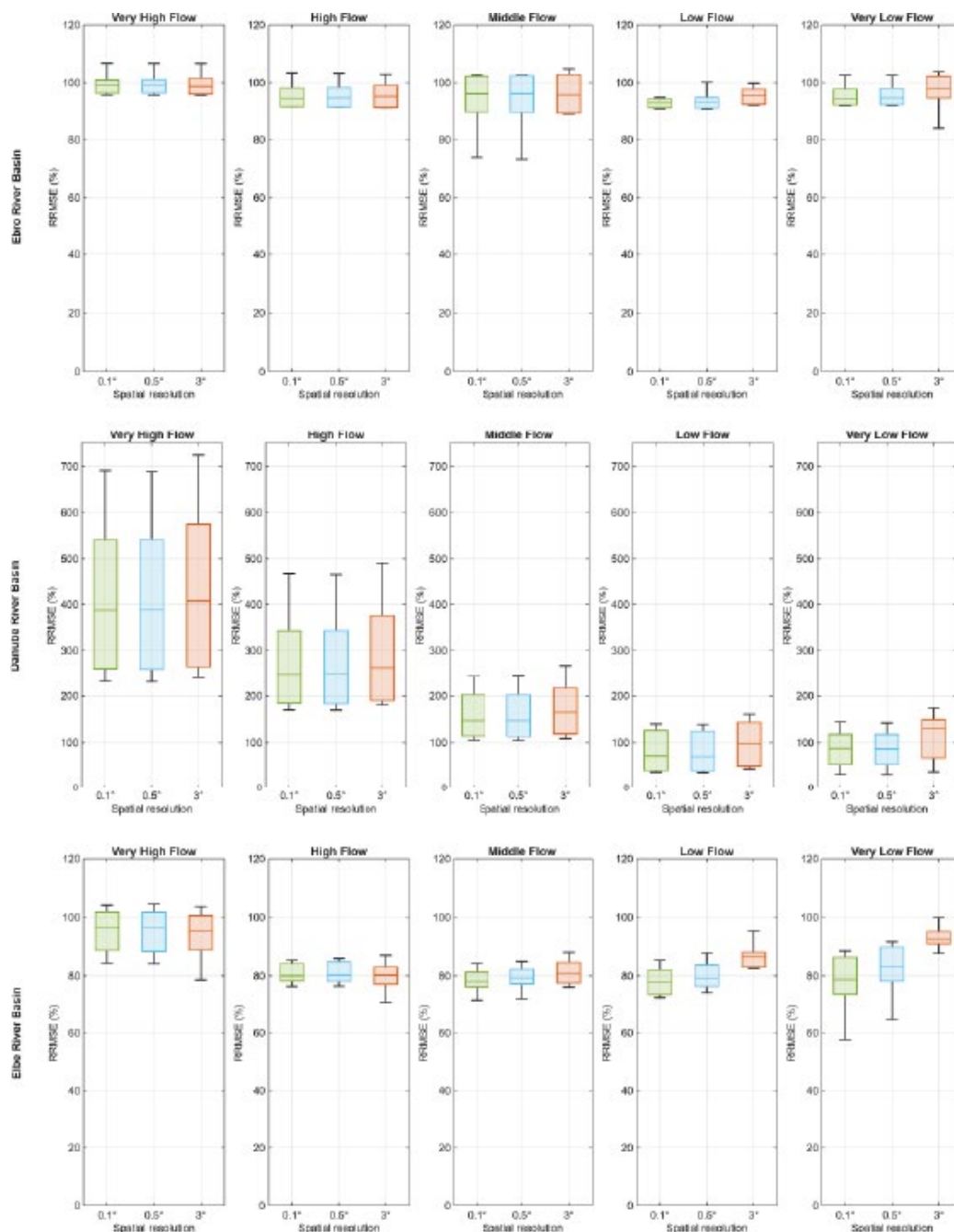


Figure 4: The figure is organized with rows representing the three pilot basins and columns corresponding to the five flow regimes analysed. Each panel shows the distribution of RRMSE (%) values across all stations within that specific basin, providing a comparison of the three spatial resolutions of the GWSA dataset.

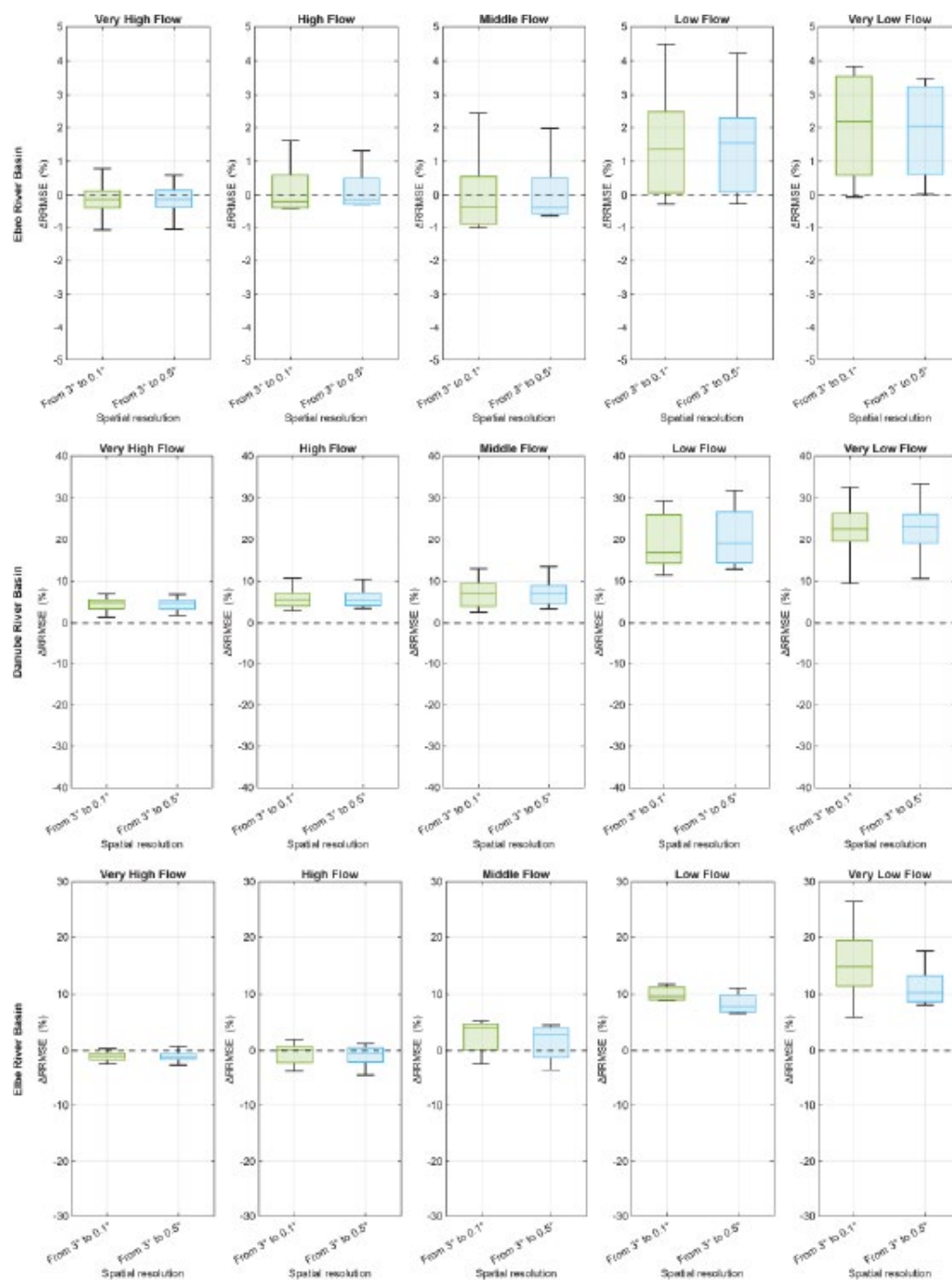


Figure 5: The figure is organized with rows representing the three pilot basins and columns corresponding to the five flow regimes analysed. Each panel shows the distribution of $\Delta RRMSE$ (%) values across all stations within that specific basin, providing a comparison of the performance variations observed when moving from the 3° spatial resolution to the finer 0.1° and 0.5° resolutions.

5.3 Response of river discharge estimation to groundwater storage resolution across baseflow contribution

The analyses presented in this section combine the information from the previous analyses by relating the RRMSE to the BF/Q ratio computed at each selected gauging station, as shown in Fig. 6. This approach allows the simultaneous assessment of the role of baseflow contribution and flow regime in controlling the sensitivity of river discharge estimates to GWSA spatial resolution. The three pilot basins are characterised by substantially different BF/Q ratios, as shown in Appendix B (Table B1). The Ebro river basin shows the lowest median BF/Q (41%), with values ranging from 35% at Caparroso to 45% at Ascò and Tortosa. The Danube river basin presents intermediate values, with a median BF/Q of 68% and a range from 62% (Schaerding) to 70% (Ceatal Izmail). The Elbe river basin exhibits the highest baseflow contribution, with a median BF/Q of 96% and values ranging from 93% (Berlin Muehlendamm and Ketzin) to 98% (Usti Nad Labem, Torgau and Aken). The three basins thus occupy distinct and non-overlapping BF/Q intervals, providing a well-defined gradient for the analysis. For the Ebro river basin, characterised by the lowest BF/Q, the RRMSE values computed across the different stations and flow regimes show limited variability, ranging from a minimum of 68% (Seros, MF, 3° configuration) to a maximum of 107% (Caparroso, VHF, 0.1° configuration). The differences between the three GWSA configurations are small across all flow regimes. For the Danube river basin, the RRMSE values exhibit a pronounced gradient across the flow regimes. The highest values are recorded for the VHF regime, reaching 726% at Sremska Mitrovica with the 3° configuration, reflecting the difficulty of the model in capturing extreme flood peaks at this station. The lowest RRMSE values are recorded for the LF and VLF regimes, with a minimum of 30% at Bazias (VLF, 0.1° configuration). The differences between the three GWSA configurations become increasingly visible in the LF and VLF regimes, where the 3° configuration systematically produces higher RRMSE values. The Elbe river basin, characterised by the highest BF/Q, shows a less extreme gradient across the flow regimes compared to the Danube, with RRMSE values ranging from 54% (Usti Nad Labem, LF, 0.1° configuration) to 104% (Torgau, VHF, 0.5° configuration). The separation between the three configurations is most visible in the LF and VLF regimes. Figure 7 shows the Δ RRMSE as a function of BF/Q for each flow regime. For the Ebro, the Δ RRMSE values are small across all regimes, with median values ranging between -1% and +2%. The only notable improvement is observed at Gelsa (BF/Q = 44%), which shows a reduction in RRMSE of approximately 8% in the VLF regime when moving from 3° to 0.1°. Conversely, Seros (BF/Q = 37%) shows a consistent increase in RRMSE across all regimes when adopting finer resolutions, with a deterioration of approximately 10% in the LF regime. For the Danube, the Δ RRMSE values show a clear gradient from the VHF to the VLF regime. The VHF regime shows only a modest improvement (1–7% across stations), while the LF and VLF regimes show the most pronounced reductions in RRMSE. The largest improvement is observed at Sremska Mitrovica (BF/Q = 67%), with a reduction of 29% in the LF and 44% in the VLF regime when moving from 3° to 0.1°. The median improvement in the VLF regime across all Danube stations is approximately 22% for the 0.1° configuration and 23% for the 0.5° configuration. All eight Danube stations show a positive Δ RRMSE across every flow regime, indicating a consistent improvement in river discharge estimation when moving from the 3° baseline to either finer configuration. For the Elbe, the Δ RRMSE values in the VHF and HF regimes are negative or close to zero for all stations, with a median deterioration of approximately 1%. From the MF regime onwards, the



Δ RRMSE becomes progressively positive. The station showing the widest range of behaviour is Berlin Muehlendamm (BF/Q = 93%), with a deterioration of 16% in the MF regime when moving from 3° to 0.5° and an improvement of 22% in the VLF regime when moving from 3° to 0.1°. The LF and VLF regimes show the most consistent improvements, with median reductions increasing systematically with BF/Q. Overall, the Δ RRMSE values increase systematically with BF/Q across all three pilot basins, with the largest reductions in RRMSE consistently observed in the LF and VLF regimes.

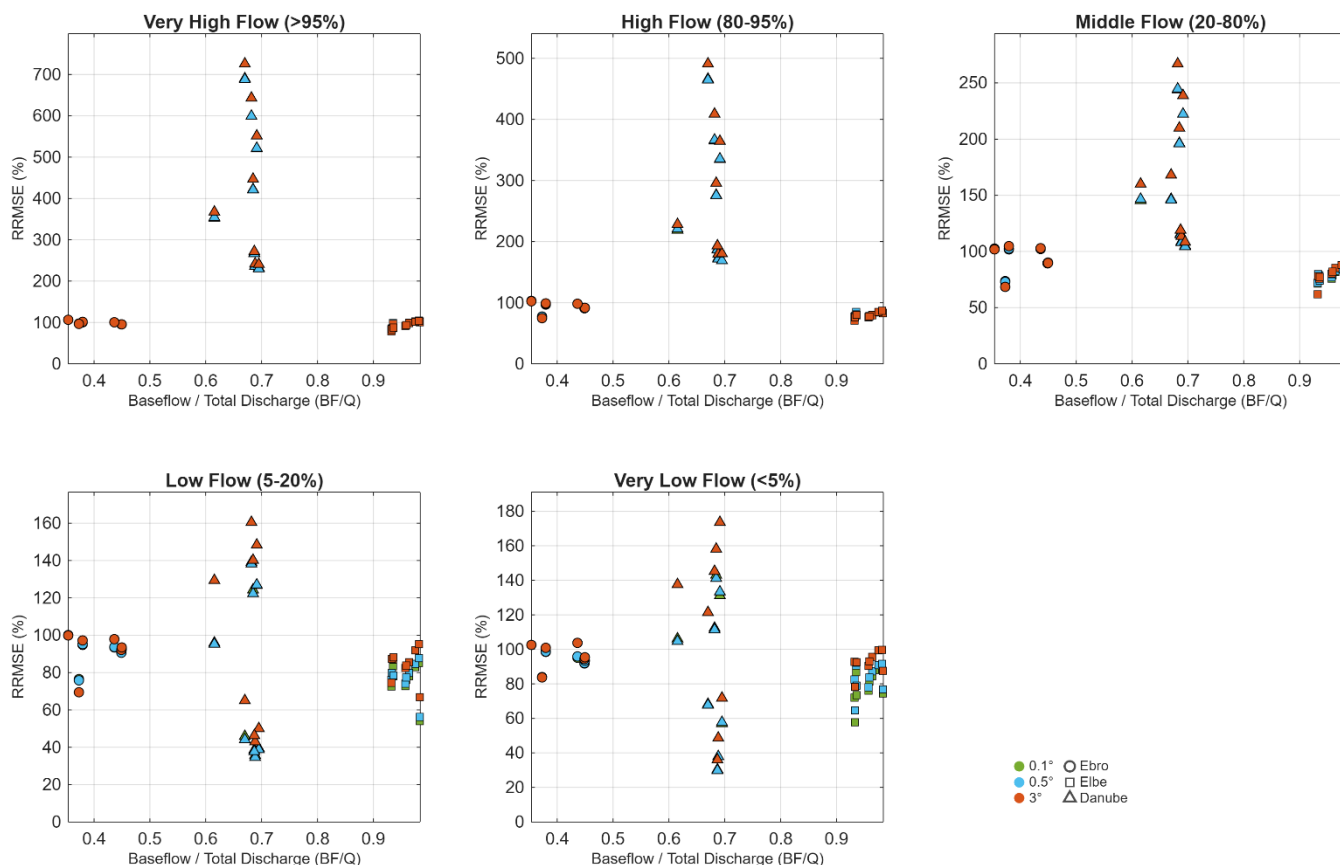
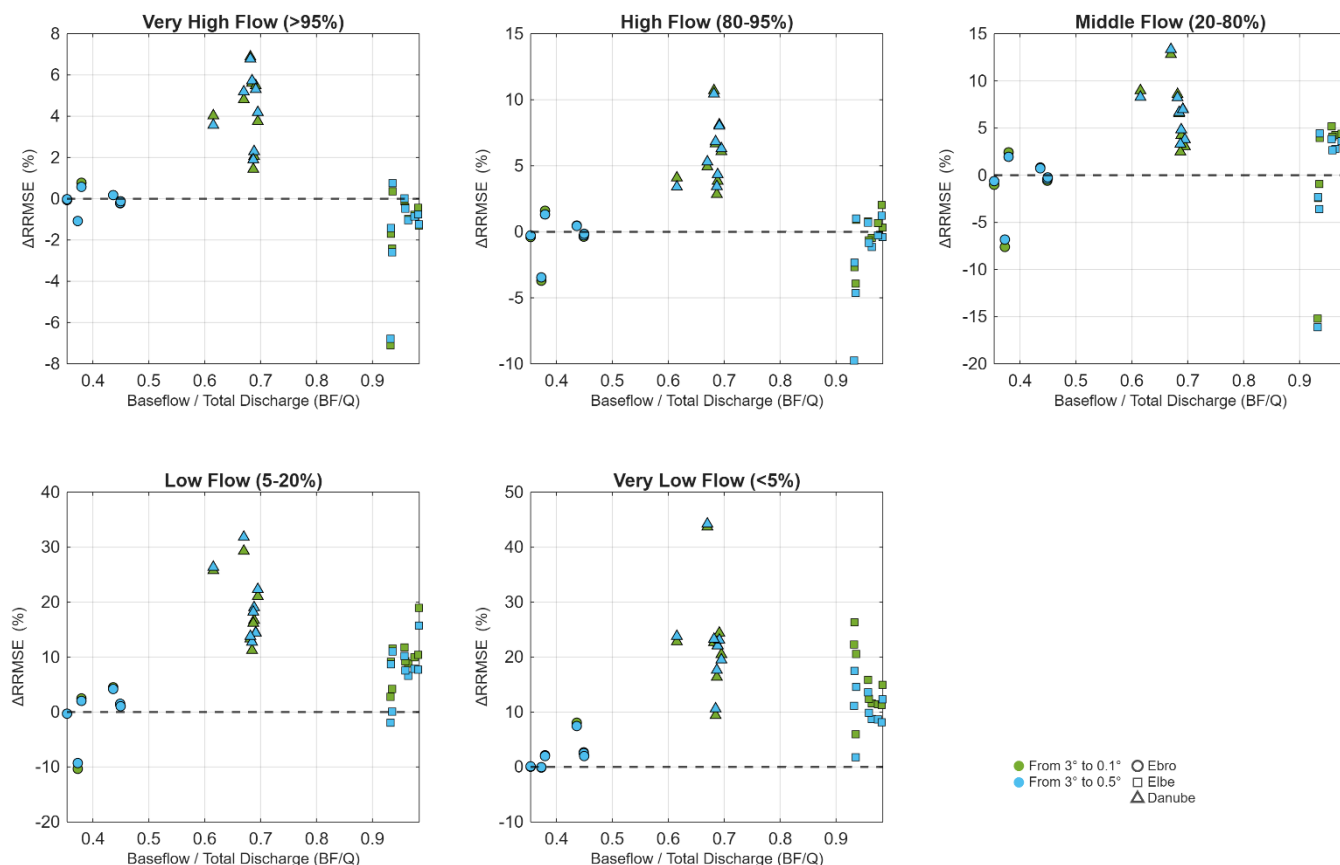


Figure 6: The figure shows the RRMSE (%) values for each flow regime, plotted against the BF/Q ratio for all analysed stations across the three pilot basins.



440 **Figure 7: The figure shows the ΔRRMSE (%) values for each flow regime, plotted against the BF/Q ratio for all analysed stations across the three pilot basins.**

6. Discussion

Although the experiments presented here rely on model-based GWSA rather than directly on GRACE TWSA, they can be interpreted within the broader framework of the GRACE literature. The modelled GWSA field is a W3RA-model state variable
 445 constrained by the assimilation of GRACE TWSA; consequently, it inherits the groundwater-relevant information contained in the gravimetry observations together with the native spatial resolution and basin-scale representativeness limitations of the assimilated TWSA product. The findings can therefore be interpreted in light of the extensive literature on the spatial resolution of GRACE-derived products. Previous studies suggested that the benefit of finer TWSA resolution should increase as basin area decreases because GRACE-derived products become progressively less representative below drainage areas of
 450 approximately 60,000–100,000 km² (Save et al., 2016; Vishwakarma et al., 2018). Our results only partially support this expectation. Smaller basins (<50,000 km²) indeed exhibit the highest errors in the 3° configuration, confirming the reduced representativeness of GRACE-derived storage estimates at these scales. However, the improvements obtained by moving from



3° to 0.5° or 0.1° remain modest and highly station dependent, with no systematic relationship between drainage area and ΔRRMSE .

455 This behaviour is likely explained by the experimental design. The three GWSA configurations do not represent independent products with progressively increasing spatial information, but different smoothing levels of the same native 0.1° dataset. Consequently, finer configurations mainly recover spatial variability already contained in the original product rather than introducing genuinely new information. Small basins therefore remain constrained by the intrinsic representativeness of the underlying GWSA estimates regardless of the smoothing applied, while for intermediate and large basins additional spatial

460 detail contributes little to basin-integrated river discharge estimates once the effective support scale of the GWSA product is exceeded. These findings indicate that drainage area is a useful criterion for assessing the applicability of GRACE-derived observations, but not for predicting where increasing GWSA spatial resolution will substantially improve river discharge simulations. More importantly, these results, obtained through synthetic experiments, suggest that the value of higher-resolution GWSA products depends less on the spatial support of the observations than on the extent to which groundwater

465 dynamics actively control streamflow generation. In other words, improving observation resolution alone is insufficient unless groundwater storage plays a dominant role in basin hydrology. This interpretation is supported by the analysis across flow regimes. For all three pilot basins, improvements associated with finer GWSA resolution are negligible during very high and high flows, whereas they become systematic during low and very low flows. Although the magnitude of the improvement differs among basins, the overall pattern is remarkably consistent. This behaviour agrees both with previous studies

470 highlighting the increasing influence of groundwater during low-flow conditions (Pfannerstill et al., 2014; Hellwig et al., 2021) and with the structure of the hydrological model adopted here. Streamflow is represented as the sum of a fast runoff component, controlled primarily by precipitation and shallow soil moisture, and a slow component driven by GWS (Camici et al., 2022; Leopardi et al., 2026). During high flows, river discharge is controlled by the fast runoff component rather than by the slow, groundwater-fed contribution, so that the simulated hydrographs are largely insensitive to GWS resolution. As flow decreases,

475 the contribution of the slow component progressively increases, amplifying the influence of GWS and making the benefit of finer spatial resolution more evident. Flow regime therefore identifies the hydrological conditions under which higher-resolution GWS information becomes relevant, but it does not explain why the magnitude of the improvement differs substantially among basins. Additional insight is provided by the baseflow-to-discharge ratio (BF/Q), which links the response to the relative importance of groundwater within the basin. Basins characterised by a larger groundwater contribution become

480 sensitive to GWSA resolution under progressively higher flow conditions. In the Elbe, where baseflow sustains most of the river discharge, improvements already emerge under medium-flow conditions. In the Danube, sensitivity becomes pronounced mainly during low flows, whereas in the Ebro, characterised by the smallest baseflow contribution, the effect of spatial resolution remains limited throughout the flow spectrum. This behaviour is consistent with previous studies showing that groundwater influence extends beyond classical low-flow conditions depending on basin storage properties (Fang and Shen,

485 2017; Berghuijs and Slater, 2023). At the same time, BF/Q does not appear to predict the magnitude of the improvement. Despite exhibiting the highest BF/Q , the Elbe shows smaller performance gains than the Danube, largely because its baseline



model performance is already comparatively high, leaving less room for improvement. The magnitude of the benefit therefore reflects not only groundwater dominance but also the initial model skill and other basin- or model-specific factors that could not be isolated in this study. Taken together, these results show that the added value of higher-resolution GWS products should not be assessed solely in terms of basin size. Instead, it emerges when groundwater dynamics are more hydrologically relevant, i.e., in LF and VLF conditions when the baseflow contribution to streamflow is dominant. This finding has important implications for future high-resolution gravimetry missions, suggesting that their greatest hydrological benefit is likely to be achieved in groundwater-dominated systems rather than uniformly across all river basins.

6.1 Limitations and challenges

The interpretation of the results presented in this study is subject to several methodological limitations. The first concerns the size and representativeness of the analysed sample. The experiments rely on three pilot basins and 24 gauging stations, unevenly distributed among the study areas, which limits the statistical robustness of the identified patterns and their generalisability beyond the selected cases. Moreover, the selected basins only cover part of the observed spectrum of baseflow contribution to total discharge. Extending the analysis to a larger and more diverse set of basins would allow the identified relationships between GWSA resolution, flow regime, and BF/Q to be tested over a broader range of hydrological conditions. Such an extension would require moving beyond the European domain, where the availability of consistent GWSA products currently constrains basin selection, and considering regions with markedly different hydrological regimes, such as arctic, tropical, or arid environments. Further limitations arise from the experimental design. As discussed in Sect. 6, the three GWSA configurations (3° , 0.5° , and 0.1°) do not represent independent observations but are generated from the same native 0.1° W3RA dataset through progressive spatial smoothing. The experiments therefore quantify the effect of degrading and recovering spatial detail within a single dataset rather than the benefit of an independent higher-resolution gravimetry product, such as those expected from future missions like NGGM-MAGIC. The improvements identified here may consequently represent a conservative estimate of the potential benefits achievable with next-generation observations. In addition, the present analysis isolates the effect of spatial resolution while keeping the temporal resolution of the GWSA input unchanged. Future gravimetry missions are expected to improve not only the spatial characteristics of terrestrial water storage observations but also their temporal sampling and the representation of storage variability. The performance gains quantified here should therefore be interpreted as the contribution of enhanced spatial resolution alone and not as the total improvement expected from future gravity missions. The model calibration strategy introduces an additional constraint. Calibrating parameters at the basin outlet and transferring them to the remaining stations within the same basin avoids compensating for differences in GWSA resolution through independent calibration, but it also limits parameter adaptability, particularly in large and heterogeneous basins such as the Danube. This contributes to the high errors observed at some stations and highlights the trade-off between isolating the effect of data resolution and preserving local model performance.

Finally, while this study identifies consistent relationships between GWSA resolution, flow regime, and baseflow contribution, the factors controlling the magnitude of the improvement remain only partially understood. No basin descriptor investigated



520 here was sufficient to quantitatively predict the expected benefit of finer-resolution GWS information, suggesting that additional hydrological and modelling factors should be considered in future studies.

7. Conclusions and outlook

This study investigated whether the benefits of increasing GWSA spatial resolution for river discharge estimation can be generalised across river basins or whether they depend on specific hydrological characteristics. Through a systematic set of synthetic spatial-resolution degradation experiments, the sensitivity of river discharge simulations to GWSA resolution was evaluated across three European pilot basins characterised by different drainage areas and baseflow contributions. The results show that basin area alone is not a sufficient predictor of the benefit associated with finer-resolution GWSA information. Although smaller basins exhibit the largest errors when using the coarsest configuration, the improvement obtained from increasing spatial resolution is limited and highly heterogeneous, with no systematic area-dependent gradient. Drainage area therefore remains a relevant descriptor of the applicability of GRACE-derived storage products, but it does not identify where enhanced spatial resolution will effectively improve river discharge simulations. A stronger and more consistent control is provided by the hydrological role of groundwater in streamflow generation. Across all basins analysed, the benefit of finer GWSA resolution emerges primarily during low and very low flow conditions, when the contribution of the slow component driven by groundwater storage dominates river discharge. Flow regime therefore represents the first-order indicator of the conditions under which higher-resolution GWSA information becomes relevant. The contribution of baseflow to total discharge (BF/Q) further refines this interpretation by controlling the position along the flow-regime spectrum at which sensitivity to GWSA resolution emerges. Basins with larger baseflow contributions become sensitive to spatial resolution under progressively higher flow conditions. Considering the full spectrum of flow regimes, ranging from VHF to VLF, the BF/Q ratio can be used as a predictor to identify from which flow regime onward the quality of the estimates is influenced by the improvement of GWSA information. However, BF/Q alone does not explain the magnitude of the improvement, which remains controlled by additional basin-specific and modelling factors. Identifying predictors capable of quantifying not only where, but also how much, finer-resolution GWS products improve river discharge simulations remains an important challenge for future research.



545

Data availability

All data used in this study are freely available at the links provided in Section 3.2, or from the corresponding author upon reasonable request.

Author contributions

550 FL, EF, LB, and SC designed the study; FL and EF provided the data; FL and SC conducted the analyses; FL and SC wrote the manuscript draft; FL, EF, JD, CS, LB, and SC reviewed and edited the manuscript.

Competing interests

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

Financial support

555 European Space Agency, ESA, funded the research through the STREAM-NEXT Project (EO Science for Society Permanently Open Call 4000126745/19/I-NB-CCN2).

Acknowledgements

The authors wish to express their sincere gratitude to the Global Runoff Data Center (GRDC) that provided most of the river discharge data.

560 During the preparation of this work, the authors used Claude (Anthropic) to improve the readability and language of the manuscript. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

565



References

- Agarwal, V., Akyilmaz, O., Shum, C.K., Feng, W., Yang, T.-Y., Forootan, E., Syed, T.H., Haritashya, U.K., Uz, M., 2023. Machine learning based downscaling of GRACE-estimated groundwater in Central Valley, California. *Science of The Total Environment* 865, 161138. <https://doi.org/10.1016/j.scitotenv.2022.161138>
- Al-Aizari, A.R., Wang, L., Gao, D., Beck, H., 2026. Evident Dependence of Dynamics of Baseflow on Groundwater Across the Contiguous United States. *Water Resources Research* 62, e2025WR042225. <https://doi.org/10.1029/2025WR042225>
- Beck, H.E., Dijk, A.I.J.M. van, Miralles, D.G., Jeu, R.A.M. de, Bruijnzeel, L.A. (Sampurno), McVicar, T.R., Schellekens, J., 2013. Global patterns in base flow index and recession based on streamflow observations from 3394 catchments. *Water Resources Research* 49, 7843–7863. <https://doi.org/10.1002/2013WR013918>
- Beck, H.E., McVicar, T.R., Vergopolan, N., Berg, A., Lutsko, N.J., Dufour, A., Zeng, Z., Jiang, X., van Dijk, A.I.J.M., Miralles, D.G., 2023. 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>
- Beck, H.E., Roo, A. de, Dijk, A.I.J.M. van, 2015. Global Maps of Streamflow Characteristics Based on Observations from Several Thousand Catchments*. *Journal of Hydrometeorology* 16, 1478–1501. <https://doi.org/10.1175/JHM-D-14-0155.1>
- Berghuijs, W.R., Slater, L.J., 2023. Groundwater shapes North American river floods. *Environ. Res. Lett.* 18, 034043. <https://doi.org/10.1088/1748-9326/acbecc>
- Bierkens, M.F.P., Bell, V.A., Burek, P., Chaney, N., Condon, L.E., David, C.H., Roo, A. de, Döll, P., Drost, N., Famiglietti, J.S., Flörke, M., Gochis, D.J., Houser, P., Hut, R., Keune, J., Kollet, S., Maxwell, R.M., Reager, J.T., Samaniego, L., Sudicky, E., Sutanudjaja, E.H., Giesen, N. van de, Winsemius, H., Wood, E.F., 2015. Hyper-resolution global hydrological modelling: what is next? *Hydrological Processes* 29, 310–320. <https://doi.org/10.1002/hyp.10391>
- Camici, S., Giuliani, G., Brocca, L., Massari, C., Tarpanelli, A., Farahani, H.H., Sneeuw, N., Restano, M., Benveniste, J., 2022. Synergy between satellite observations of soil moisture and water storage anomalies for runoff estimation. *Geoscientific Model Development* 15, 6935–6956. <https://doi.org/10.5194/gmd-15-6935-2022>
- Cao, F., Gerges, E., Vargas, C.R., Pearson, A.R., Lictevout, E., 2026. In-situ groundwater level monitoring reveals worldwide status, trends and impacts. *Environ. Res. Lett.* 21, 103002. <https://doi.org/10.1088/1748-9326/ae651f>
- Cui, L., Li, Y., Zhong, B., An, J., Meng, J., Guo, H., Xu, C., 2025. Assessing the impact of 2022 extreme drought on the Yangtze River basin using downscaled GRACE/GRACE-FO data obtained by partitioned random forest algorithm. *International Journal of Remote Sensing* 46, 1219–1247. <https://doi.org/10.1080/01431161.2024.2427915>
- Cuthbert, M.O., Gleeson, T., Bierkens, M.F.P., Ferguson, G., Taylor, R.G., 2024. Concerns regarding proposed groundwater Earth system boundary. *Nature* 635, E4–E5. <https://doi.org/10.1038/s41586-024-08082-9>



- Daras, I., March, G., Pail, R., Hughes, C.W., Braitenberg, C., Güntner, A., Eicker, A., Wouters, B., Heller-Kaikov, B., Pivetta, T., Pastorutti, A., 2024. 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>
- 600 Dari, J., Filippucci, P., Brocca, L., Quast, R., Vreugdenhil, M., Miralles, D.G., Morbidelli, R., Saltalippi, C., Flammini, A., 2025. A novel approach for estimating groundwater recharge leveraging high-resolution satellite soil moisture. *Journal of Hydrology* 652, 132678. <https://doi.org/10.1016/j.jhydrol.2025.132678>
- Dembélé, M., Schaefli, B., van de Giesen, N., Mariéthoz, G., 2020. Suitability of 17 gridded rainfall and temperature datasets for large-scale hydrological modelling in West Africa. *Hydrology and Earth System Sciences* 24, 5379–5406. <https://doi.org/10.5194/hess-24-5379-2020>
- 605 Döll, P., Kaspar, F., Lehner, B., 2003. A global hydrological model for deriving water availability indicators: model tuning and validation. *Journal of Hydrology* 270, 105–134. [https://doi.org/10.1016/S0022-1694\(02\)00283-4](https://doi.org/10.1016/S0022-1694(02)00283-4)
- Fang, K., Shen, C., 2017. Full-flow-regime storage-streamflow correlation patterns provide insights into hydrologic functioning over the continental US. *Water Resources Research* 53, 8064–8083. <https://doi.org/10.1002/2016WR020283>
- 610 Forootan, E., Mehrnegar, N., Schumacher, M., Schiettekatte, L.A.R., Jagdhuber, T., Farzaneh, S., van Dijk, A.I.J.M., Shamsudduha, M., Shum, C.K., 2024. Global groundwater droughts are more severe than they appear in hydrological models: An investigation through a Bayesian merging of GRACE and GRACE-FO data with a water balance model. *Science of The Total Environment* 912, 169476. <https://doi.org/10.1016/j.scitotenv.2023.169476>
- Gleeson, T., Wada, Y., Bierkens, M.F.P., van Beek, L.P.H., 2012. Water balance of global aquifers revealed by groundwater footprint. *Nature* 488, 197–200. <https://doi.org/10.1038/nature11295>
- 615 Gupta, H.V., Kling, H., Yilmaz, K.K., Martinez, G.F., 2009. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology* 377, 80–91. <https://doi.org/10.1016/j.jhydrol.2009.08.003>
- Heinicke, S., Volkholz, J., Schewe, J., Gosling, S.N., Müller Schmied, H., Zimmermann, S., Mengel, M., Sauer, I.J., Burek, P., Chang, J., Kou-Giesbrecht, S., Grillakis, M., Guillaumot, L., Hanasaki, N., Koutroulis, A., Otta, K., Qi, W., Satoh, Y., Stacke, T., Yokohata, T., Frieler, K., 2024. Global hydrological models continue to overestimate river discharge. *Environ. Res. Lett.* 19, 074005. <https://doi.org/10.1088/1748-9326/ad52b0>
- Hellwig, J., Stoezle, M., Stahl, K., 2021. Groundwater and baseflow drought responses to synthetic recharge stress tests. *Hydrology and Earth System Sciences* 25, 1053–1068. <https://doi.org/10.5194/hess-25-1053-2021>
- 625 Hu, Y., Cai, X., Xu, Y.-P., Gu, H., Xie, J., Dai, S., 2026. A semi-supervised LSTM framework for spatiotemporal downscaling of GRACE-derived terrestrial water storage anomalies to improve flood monitoring in the Yarlung Tsangpo River basin. *Journal of Hydrology* 671, 135275. <https://doi.org/10.1016/j.jhydrol.2026.135275>
- Huffman, G.J., Bolvin, D.T., Braithwaite, D., Hsu, K.-L., Joyce, R.J., Kidd, C., Nelkin, E.J., Sorooshian, S., Stocker, E.F., Tan, J., Wolff, D.B., Xie, P., 2020. Integrated Multi-satellite Retrievals for the Global Precipitation Measurement (GPM) Mission (IMERG), in: Levizzani, V., Kidd, C., Kirschbaum, D.B., Kummerow, C.D., Nakamura, K., Turk, F.J. (Eds.), *Satellite*
- 630



- Precipitation Measurement: Volume 1. Springer International Publishing, Cham, pp. 343–353. https://doi.org/10.1007/978-3-030-24568-9_19
- Jasechko, S., Seybold, H., Perrone, D., Fan, Y., Shamsudduha, M., Taylor, R.G., Fallatah, O., Kirchner, J.W., 2024. Rapid groundwater decline and some cases of recovery in aquifers globally. *Nature* 625, 715–721. <https://doi.org/10.1038/s41586-023-06879-8>
- Landerer, F.W., Swenson, S.C., 2012. Accuracy of scaled GRACE terrestrial water storage estimates. *Water Resources Research* 48. <https://doi.org/10.1029/2011WR011453>
- Leopardi, F., Dari, J., Saltalippi, C., Brocca, L., Volden, E., Fernández-Prieto, D., Saemian, P., Sneeuw, N., Tourian, M.J., Camici, S., 2026. A satellite-based approach for estimating runoff and river discharge in the Pan-Arctic region from 2003 to 2022. *Remote Sensing of Environment* 338, 115353. <https://doi.org/10.1016/j.rse.2026.115353>
- Longuevergne, L., Scanlon, B.R., Wilson, C.R., 2010. GRACE Hydrological estimates for small basins: Evaluating processing approaches on the High Plains Aquifer, USA. *Water Resources Research* 46. <https://doi.org/10.1029/2009WR008564>
- Mehrnagar, N., Brocca, L., Forootan, E., 2026. Groundwater storage changes within europe: a constrained bayesian integration of GRACE(-FO) data into the W3RA model. <https://doi.org/10.6084/m9.figshare.31073527.v2>
- Mehrnagar, N., Jones, O., Singer, M.B., Schumacher, M., Jagdhuber, T., Scanlon, B.R., Rateb, A., Forootan, E., 2021. Exploring groundwater and soil water storage changes across the CONUS at 12.5 km resolution by a Bayesian integration of GRACE data into W3RA. *Science of The Total Environment* 758, 143579. <https://doi.org/10.1016/j.scitotenv.2020.143579>
- Mehrnagar, N., Schumacher, M., Jagdhuber, T., Forootan, E., 2023. Making the Best Use of GRACE, GRACE-FO and SMAP Data Through a Constrained Bayesian Data-Model Integration. *Water Resources Research* 59, e2023WR034544. <https://doi.org/10.1029/2023WR034544>
- Middendorf, K., Jensen, L., Schlaak, M., Haas, J., Dobslaw, H., Pail, R., Güntner, A., Eicker, A., 2026. Benefits of future satellite gravimetry missions for characterizing extreme wet events in terrestrial water storage. *Advances in Space Research*. <https://doi.org/10.1016/j.asr.2026.07.075>
- Mikaeeli Ahamadabad, F., Dastjerdi, P.A., Nasser, M., 2025. Spatiotemporal GRACE TWS downscaling using statistical and machine learning methods: Random Forest, area-to-area kriging, and hybrid methods. *Journal of Hydrology: Regional Studies* 62, 102885. <https://doi.org/10.1016/j.ejrh.2025.102885>
- Mohanasundaram, S., Mekonnen, M.M., Haacker, E., Ray, C., Lim, S., Shrestha, S., 2021. An application of GRACE mission datasets for streamflow and baseflow estimation in the Conterminous United States basins. *Journal of Hydrology* 601, 126622. <https://doi.org/10.1016/j.jhydrol.2021.126622>
- 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., Thépaut, J.-N., 2021. ERA5-Land: a state-of-the-art global reanalysis dataset for land applications. *Earth System Science Data* 13, 4349–4383. <https://doi.org/10.5194/essd-13-4349-2021>



- Petersen-Perlman, J.D., Aguilar-Barajas, I., Megdal, S.B., 2022. Drought and groundwater management: Interconnections, challenges, and policy responses. *Current Opinion in Environmental Science & Health* 28, 100364. <https://doi.org/10.1016/j.coesh.2022.100364>
- Pfannerstill, M., Guse, B., Fohrer, N., 2014. Smart low flow signature metrics for an improved overall performance evaluation of hydrological models. *Journal of Hydrology* 510, 447–458. <https://doi.org/10.1016/j.jhydrol.2013.12.044>
- Pfeffer, J., Meyssignac, B., Bingham, R., Blazquez, A., Bouih, M., Braitenberg, C., Brocca, L., Dobsław, H., Ferrari, R., Forootan, E., Gerdener, H., Javed, M.T., Jensen, L., Klemann, V., Kremer, A., Kusche, J., Larnicol, G., Liaqat, M.U., Leopardi, F., Mamagiannou, E.-M., Maurizio, G., Pail, R., Panet, I., Papanikolaou, T., Saemian, P., Sasgen, I., Schumacher, M., Schlaak, M., Shihora, L., Sobouti, A., Sneeuw, N., Tourian, M.J., Tsoulis, D., Vergos, G., Wouters, B., Yang, F., Daras, I., 2026. Impact of the NGGM and MAGIC future satellite gravity missions for scientific applications and operational services. <https://doi.org/10.5194/egusphere-2026-4642>
- Pradhan, R.K., Markonis, Y., Vargas Godoy, M.R., Villalba-Pradas, A., Andreadis, K.M., Nikolopoulos, E.I., Papalexiou, S.M., Rahim, A., Tapiador, F.J., Hanel, M., 2022. Review of GPM IMERG performance: A global perspective. *Remote Sensing of Environment* 268, 112754. <https://doi.org/10.1016/j.rse.2021.112754>
- Rakovec, O., Kumar, R., Attinger, S., Samaniego, L., 2016. Improving the realism of hydrologic model functioning through multivariate parameter estimation. *Water Resources Research* 52, 7779–7792. <https://doi.org/10.1002/2016WR019430>
- Retegui-Schiettekatte, L., Schumacher, M., Madsen, H., Forootan, E., 2025. Assessing daily GRACE Data Assimilation during flood events of the Brahmaputra River Basin. *Science of The Total Environment* 975, 179181. <https://doi.org/10.1016/j.scitotenv.2025.179181>
- Rodell, M., Velicogna, I., Famiglietti, J.S., 2009. Satellite-based estimates of groundwater depletion in India. *Nature* 460, 999–1002. <https://doi.org/10.1038/nature08238>
- Sadeghi, M., Tuller, M., Warrick, A.W., Babaeian, E., Parajuli, K., Gohardoust, M.R., Jones, S.B., 2019. An analytical model for estimation of land surface net water flux from near-surface soil moisture observations. *Journal of Hydrology* 570, 26–37. <https://doi.org/10.1016/j.jhydrol.2018.12.038>
- Save, H., Bettadpur, S., Tapley, B.D., 2016. High-resolution CSR GRACE RL05 mascons. *Journal of Geophysical Research: Solid Earth* 121, 7547–7569. <https://doi.org/10.1002/2016JB013007>
- Sawicz, K., Wagener, T., Sivapalan, M., Troch, P.A., Carrillo, G., 2011. Catchment classification: empirical analysis of hydrologic similarity based on catchment function in the eastern USA. *Hydrology and Earth System Sciences* 15, 2895–2911. <https://doi.org/10.5194/hess-15-2895-2011>
- Scanlon, B.R., Zhang, Z., Rateb, A., Sun, A., Wiese, D., Save, H., Beaudoin, H., Lo, M.H., Müller-Schmied, H., Döll, P., van Beek, R., Swenson, S., Lawrence, D., Croteau, M., Reedy, R.C., 2019. Tracking Seasonal Fluctuations in Land Water Storage Using Global Models and GRACE Satellites. *Geophysical Research Letters* 46, 5254–5264. <https://doi.org/10.1029/2018GL081836>



- Schumacher, M., Forootan, E., van Dijk, A.I.J.M., Müller Schmied, H., Crosbie, R.S., Kusche, J., Döll, P., 2018. Improving drought simulations within the Murray-Darling Basin by combined calibration/assimilation of GRACE data into the WaterGAP Global Hydrology Model. *Remote Sensing of Environment* 204, 212–228.
 700 <https://doi.org/10.1016/j.rse.2017.10.029>
- Schumacher, M., van Dijk, A.I.J.M., Retegui-Schiettekatte, L., Yang, F., Forootan, E., 2025. Space-based natural and human-induced water storage change quantification. *Sci Rep* 15, 18484. <https://doi.org/10.1038/s41598-025-01938-8>
- Shuai, P., Chen, X., Mital, U., Coon, E.T., Dwivedi, D., 2022. The effects of spatial and temporal resolution of gridded meteorological forcing on watershed hydrological responses. *Hydrology and Earth System Sciences* 26, 2245–2276.
 705 <https://doi.org/10.5194/hess-26-2245-2022>
- Siebert, S., Burke, J., Faures, J.M., Frenken, K., Hoogeveen, J., Döll, P., Portmann, F.T., 2010. Groundwater use for irrigation – a global inventory. *Hydrology and Earth System Sciences* 14, 1863–1880. <https://doi.org/10.5194/hess-14-1863-2010>
- Smakhtin, V.U., 2001. Low flow hydrology: a review. *Journal of Hydrology* 240, 147–186. [https://doi.org/10.1016/S0022-1694\(00\)00340-1](https://doi.org/10.1016/S0022-1694(00)00340-1)
- 710 Springer, A., De Lannoy, G., Rodell, M., Ewerdwalbesloh, Y., Gerdener, H., Khaki, M., Li, B., Li, F., Schumacher, M., Tangdamrongsub, N., Tourian, M.J., Nie, W., Kusche, J., 2026. A review of current best practices and future directions in assimilating GRACE/-FO terrestrial water storage data into numerical models. *Hydrology and Earth System Sciences* 30, 985–1022. <https://doi.org/10.5194/hess-30-985-2026>
- Sun, A.Y., Save, H., Rateb, A., Jiang, P., Scanlon, B.R., 2024. Deciphering the Role of Total Water Storage Anomalies in
 715 Mediating Regional Flooding. *Geophysical Research Letters* 51, e2023GL108126. <https://doi.org/10.1029/2023GL108126>
- Tourian, M.J., Reager, J.T., Sneeuw, N., 2018. The Total Drainable Water Storage of the Amazon River Basin: A First Estimate Using GRACE. *Water Resources Research* 54, 3290–3312. <https://doi.org/10.1029/2017WR021674>
- Tourian, M.J., Saemian, P., Ferreira, V.G., Sneeuw, N., Frappart, F., Papa, F., 2023. A copula-supported Bayesian framework for spatial downscaling of GRACE-derived terrestrial water storage flux. *Remote Sensing of Environment* 295, 113685.
 720 <https://doi.org/10.1016/j.rse.2023.113685>
- Vishwakarma, B.D., Devaraju, B., Sneeuw, N., 2018. What Is the Spatial Resolution of grace Satellite Products for Hydrology? *Remote Sensing* 10, 852. <https://doi.org/10.3390/rs10060852>
- Wagner, W., Lemoine, G., Rott, H., 1999. A Method for Estimating Soil Moisture from ERS Scatterometer and Soil Data. *Remote Sensing of Environment* 70, 191–207. [https://doi.org/10.1016/S0034-4257\(99\)00036-X](https://doi.org/10.1016/S0034-4257(99)00036-X)
- 725 Watkins, M.M., Wiese, D.N., Yuan, D.-N., Boening, C., Landerer, F.W., 2015. Improved methods for observing Earth’s time variable mass distribution with GRACE using spherical cap mascons. *Journal of Geophysical Research: Solid Earth* 120, 2648–2671. <https://doi.org/10.1002/2014JB011547>
- Xie, J., Xu, Y.-P., Yu, H., Huang, Y., Guo, Y., 2022. Monitoring the extreme flood events in the Yangtze River basin based on GRACE and GRACE-FO satellite data. *Hydrology and Earth System Sciences* 26, 5933–5954.
 730 <https://doi.org/10.5194/hess-26-5933-2022>



Xiong, J., Wang, Z., Guo, S., Wu, X., Yin, J., Wang, J., Lai, C., Gong, Q., 2022. High effectiveness of GRACE data in daily-scale flood modeling: case study in the Xijiang River Basin, China. *Nat Hazards* 113, 507–526. <https://doi.org/10.1007/s11069-022-05312-z>

735 Zhang, D., Liu, X., Bai, P., 2019. Assessment of hydrological drought and its recovery time for eight tributaries of the Yangtze River (China) based on downscaled GRACE data. *Journal of Hydrology* 568, 592–603. <https://doi.org/10.1016/j.jhydrol.2018.11.030>

Zingerle, P., Gruber, T., Pail, R., Daras, I., 2024. Constellation design and performance of future quantum satellite gravity missions. *Earth Planets Space* 76, 101. <https://doi.org/10.1186/s40623-024-02034-3>



740

Appendix A

STREAM is a conceptual, semi-distributed hydrological model that combines satellite observations of precipitation, soil moisture (SM), and snow cover fraction (SCF) with model-based air temperature and terrestrial water storage anomalies (TWSA) data. The model provides daily gridded estimates of runoff and river discharge. Its formulation assumes that satellite-
 745 derived SWI and terrestrial water storage observations can serve as proxies for soil water content at different depths, thereby avoiding the explicit representation of complex vertical surface and subsurface fluxes such as infiltration, percolation, and evapotranspiration. This results in a parsimonious formulation compared with physically based hydrological models, with lower computational and calibration costs but reduced process-level physical interpretability.

The STREAM model was originally introduced by Camici et al. (2022) and subsequently refined by Leopardi et al. (2026) to
 750 extend its applicability to river basins in the Pan-Arctic region. In both formulations, daily river discharge is represented as the sum of a fast component (R_q), primarily driven by precipitation and superficial soil moisture (here directly substituted by SWI), and a slow component (R_s), primarily driven by TWSA according to equations (1) and (2), depending on whether the snow cover fraction (SCF) exceeds the 50% threshold within the pixel:

$$R_q(t) = r(t)SWI^\alpha; R_s(t) = \beta (TWSA(t))^m \quad \text{if } SCF(t) \leq 50\% \quad (1)$$

$$755 \quad R_q(t) = 0; \quad R_s(t) = \beta [(1 - \eta SCF(t)) TWSA^*(t)]^m \quad \text{if } SCF(t) > 50\% \quad (2)$$

In the current formulation, the original variable $TWSA^*$, representing the normalised anomalies of Terrestrial Water Storage measured by GRACE and its successor GRACE-FO, has been replaced with the GWSA data. The parameters β , m and η are determined during the model calibration procedure. The first two account for the strong non-linearity linking baseflow and TWSA, while η is a damping factor that ensures the persistence of the baseflow contribution even when the soil is completely
 760 covered by snow or ice. Furthermore, in equation (2) the term $(1 - \eta SCF(t))$ acts as a filter, extracting the liquid component of the water mass to account for the subsurface fluxes feeding the baseflow. This formulation enables the continuous monitoring of baseflow even during the winter season, when the surface layer is completely frozen or snow-covered — a feature of particular relevance for the present study, as two of the selected basins (the Danube and the Elbe) are affected by significant seasonal snow cover. For a more rigorous and complete description of the model formulation, the reader is referred
 765 to Camici et al. (2022) for the original version and to Leopardi et al. (2026) for the recent extension to the Pan-Arctic region.



Appendix B

Table B1. BF/Q ratio computed at each selected gauging station in the Ebro, Danube and Elbe river basins.

#	Station name	River Basin	BF/Q
1	Caparroso	Ebro	35%
2	Seros	Ebro	37%
3	Castejon	Ebro	38%
4	Gelsa	Ebro	44%
5	Ascò	Ebro	45%
6	Tortosa	Ebro	45%
7	Schaerding	Danube	62%
8	Sremska Mitrovica	Danube	67%
9	Nagymaros	Danube	68%
10	Bogojevo	Danube	68%
11	Bazias	Danube	69%
12	Zimnicea	Danube	70%
13	Mohacs	Danube	69%
14	Ceatal Izmail	Danube	70%
15	Berlin-Mühlendamm Op	Elbe	93%
16	Ketzin	Elbe	93%
17	Calbe Grizehne	Elbe	94%
18	Havelberg Stadt	Elbe	94%
19	Ústí nad Labem	Elbe	98%
20	Torgau	Elbe	98%
21	Aken	Elbe	98%
22	Tangermünde	Elbe	96%
23	Wittenberge	Elbe	96%



24	Neu Darchau	Elbe	96%
----	-------------	------	-----

770



Appendix C

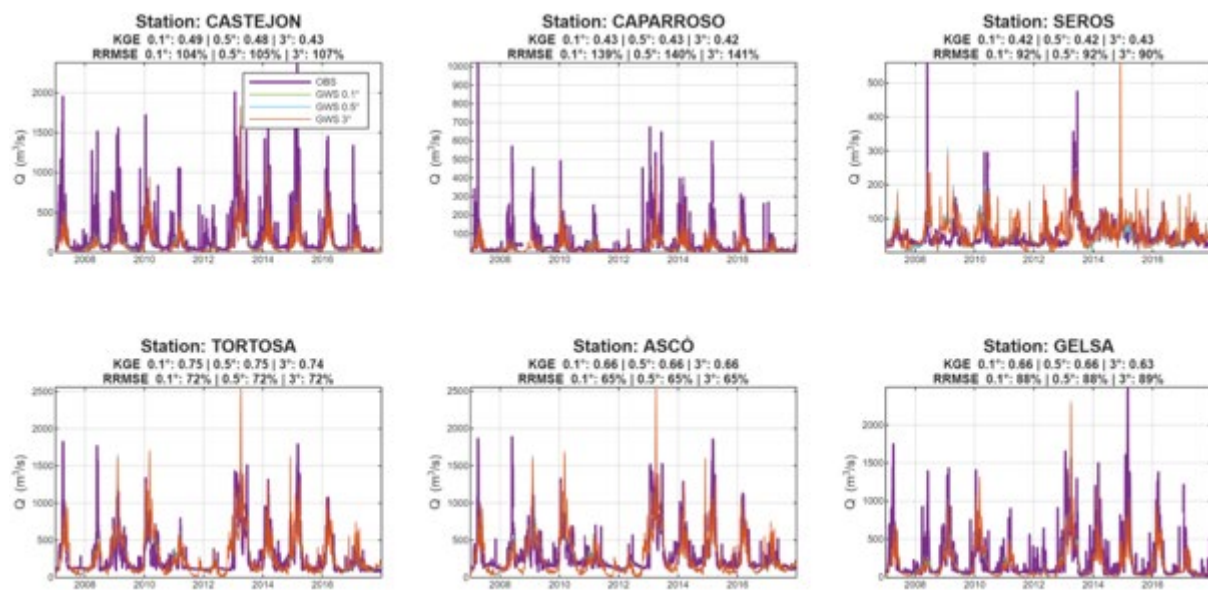
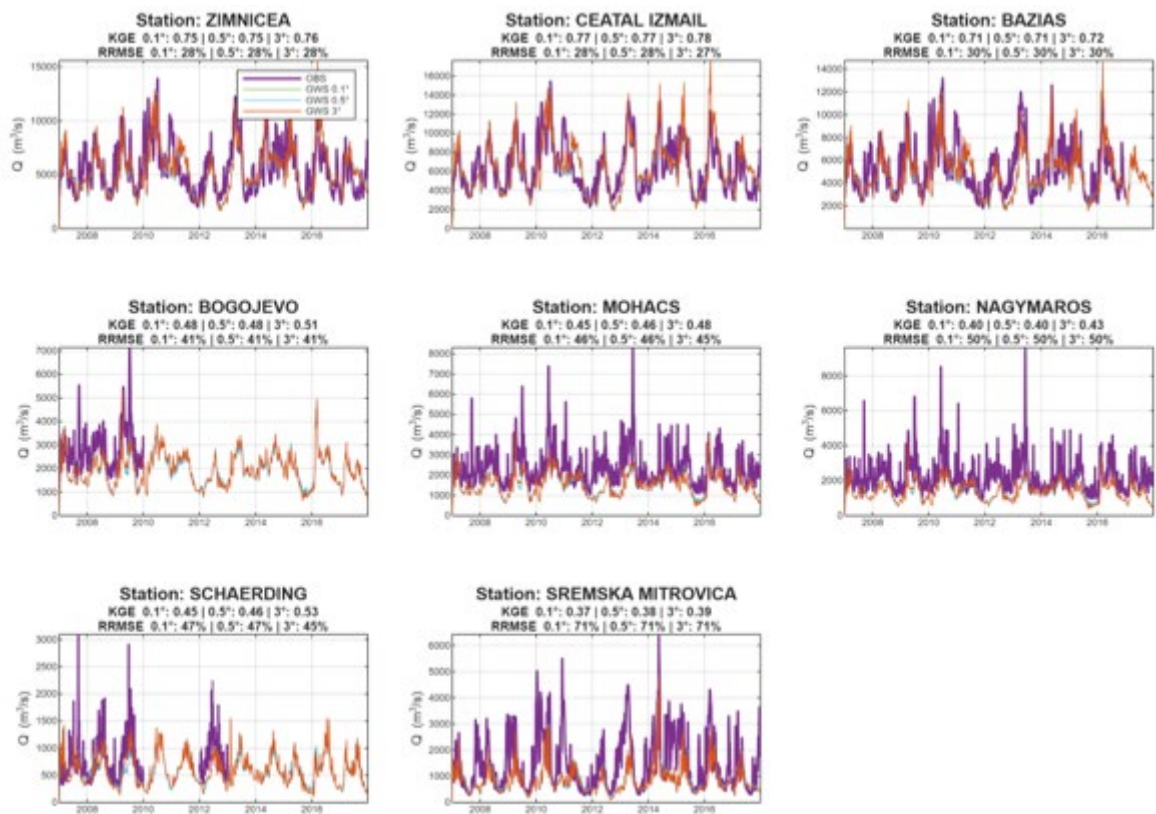
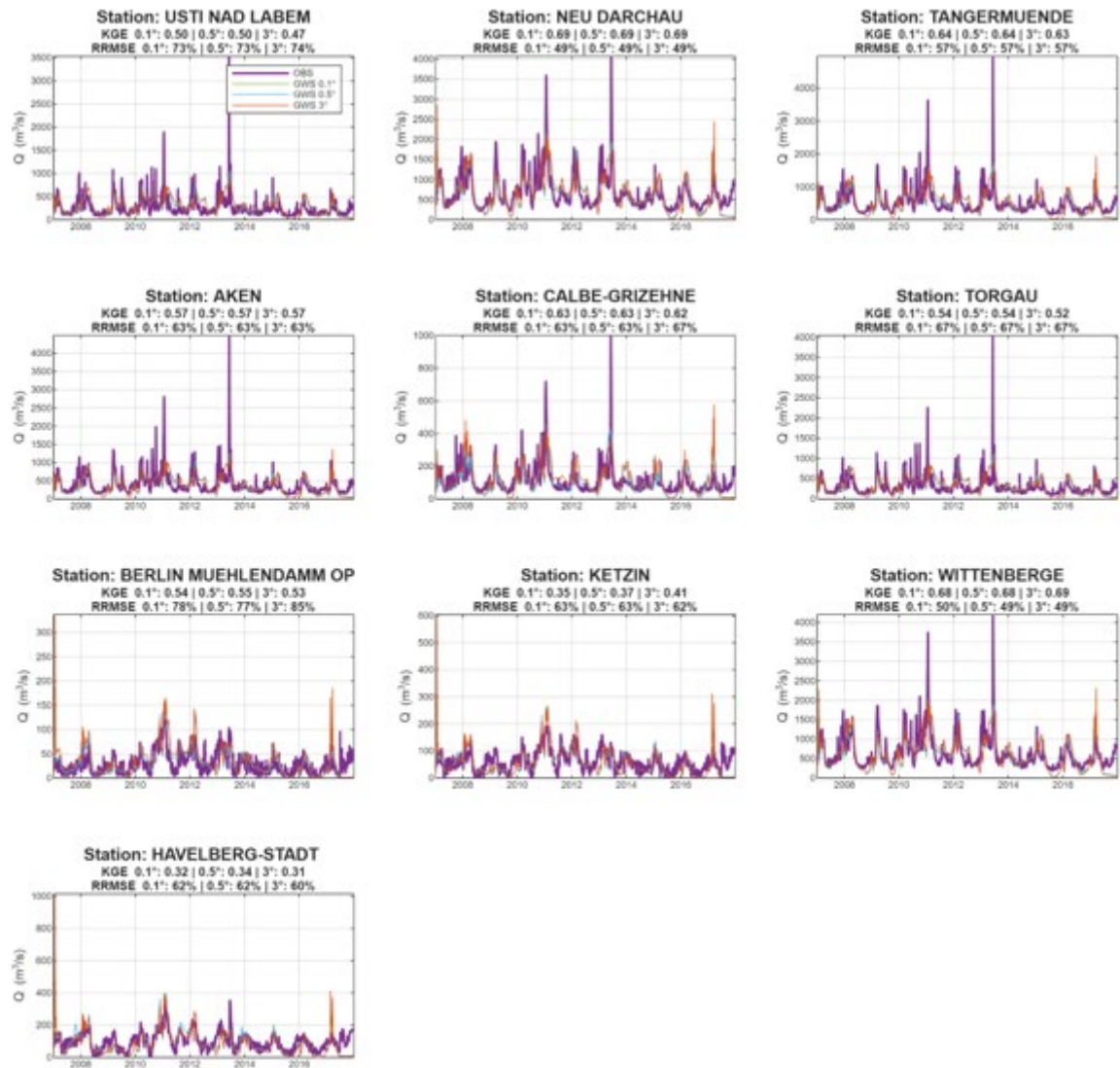


Figure C1: Comparison of observed and simulated daily river discharge for the three GWSA spatial resolutions at the selected gauging stations in the Ebro River basin (2007–2017).



775

Figure C2: Comparison of observed and simulated daily river discharge for the three GWSA spatial resolutions at the selected gauging stations in the Danube River basin (2007–2017).



780 **Figure C3.** Comparison of observed and simulated daily river discharge for the three GWSA spatial resolutions at the selected gauging stations in the Elbe River basin (2007–2017).