submitted to *Geophys. J. Int.*

Mass-change And Geosciences International Constellation (MAGIC) expected impact on science and applications

I. Daras[1⋆], G. March[2], R. Pail[3], C. W. Hughes[4], C. Braitenberg[5], A. Güntner[6,7], A. Eicker[8], B. Wouters[9], B. Heller-Kaikov[3], T. Pivetta[5], A. Pastorutti[5]

[1] *European Space Agency, Earth & Mission Science Division, ESTEC, Noordwijk, The Netherlands.*

[2] *RHEA for European Space Agency, Earth & Mission Science Division, ESTEC, Noordwijk, The Netherlands.*

[3] *Institute of Astronomical and Physical Geodesy, Technical University of Munich (TUM), Arcisstraße 21, 80333 Münch*

[4] *School of Environmental Sciences, University of Liverpool, Liverpool, U.K.,*

and National Oceanography Centre, Liverpool, U.K.

[5] *Department of Mathematics and Geosciences, University of Trieste, Via Weiss 1, 34128 Trieste, Italy.*

[6] *Helmholtz Centre Potsdam GFZ German Research Centre for Geosciences, Potsdam, Germany*

[7] *University of Potsdam, Institute of Environmental Sciences and Geography, Potsdam, Germany*

[8] *HafenCity University Hamburg, Hamburg, Germany*

[9] *Technical University Delft, Department of Geoscience and Remote Sensing, Delft, Netherlands*

Submitted on: 2023 March 18

SUMMARY

The joint ESA/NASA Mass-change And Geosciences International Constellation (MAGIC) has the objective to extend time series from previous gravity missions, including an improvement of accuracy and spatio-temporal resolution. The long-term monitoring of Earth's gravity field carries information on mass-change induced by water cycle, climate change, and mass transport processes between atmosphere, cryosphere, oceans and solid Earth. MAGIC will be composed of two satellite pairs flying in different orbit planes. The NASA/DLR–led first pair (P1) is expected to be in a near-polar orbit around 500 km of altitude; while the second

ESA–led pair (P2) is expected to be in an inclined orbit of 65–70 degrees at approximately 400 km altitude. The ESA–led pair P2 Next Generation Gravity Mission (NGGM) shall be launched after P1 in a staggered manner to form the MAGIC constellation. The addition of an inclined pair shall lead to reduction of temporal aliasing effects and consequently of reliance on de-aliasing models and post-processing. The main novelty of the MAGIC constellation is the delivery of mass-change products at higher spatial resolution, temporal (i.e. sub–weekly) resolution, shorter latency, and higher accuracy than GRACE and GRACE-FO. This will pave the way to new science applications and operational services. In this article, an overview of various fields of science and service applications for hydrology, cryosphere, oceanography, solid Earth, climate change and geodesy is provided. These thematic fields and newly enabled applications and services were analysed in the frame of the initial ESA Science Support activities for MAGIC. The analyses of MAGIC scenarios for different application areas in the field of geosciences confirmed that the double-pair configuration will significantly enlarge the number of observable mass-change phenomena by resolving smaller spatial scales with an uncertainty that satisfies evolved user requirements expressed by international bodies such as IUGG. The required uncertainty levels of dedicated thematic fields met by MAGIC unfiltered Level-2 products will benefit hydrological applications by recovering more than 90% of the major river basins worldwide at 260 km spatial resolution, cryosphere applications by enabling mass change signal separation in the interior of Greenland from those in the coastal zones and by resolving small-scale mass variability in challenging regions such as the Antarctic Peninsula, oceanography applications by monitoring meridional overturning circulation changes on time scales of years and decades, climate applications by detecting amplitude and phase changes of Terrestrial Water Storage (TWS) after 30 years in 64% and 56% of the global land areas and solid Earth applications by lowering the Earthquake detection threshold from magnitude 8.8 to magnitude 7.4 with spatial resolution increased to 333 km.

Key words: Satellite gravity, Time variable gravity, Hydrology, Global change from geodesy, Earthquake dynamics, Glaciology.

⋆ ilias.daras@esa.int

1 INTRODUCTION

Continuity and evolution of gravity missions to observe mass-change is foreseen around the turn of this decade in order to meet priority user needs not addressed by the existing and planned satellite infrastructure. The global scientific user community has been continuously addressing the need for enhanced and sustained mass change monitoring from space, via bodies such as the Global Geodetic Observing System (GGOS), International Association of Geodesy (IAG) and the International Union of Geodesy and Geophysics (IUGG) (IUGG 2015 Resolution no. 2, 2015), (IUGG 2023 Resolution no. 2, 2023). Predecessor gravity missions like CHAMP (Reigber et al. 2002), GRACE (Tapley et al. 2004), GOCE (Drinkwater et al. 2003) and GRACE Follow-On (Landerer et al. 2020) have revolutionize our understanding of the global static and temporal varying gravity field and the related monitoring of mass transport processes, but also showed limitation regarding the achievable spatial and temporal resolution of resulting gravity field product, e.g., (Thomas et al. 2017), (Rodell et al. 2018), (Cazenave et al., 2019), (Wouters et al. 2019), (Ciracì et al. 2020).

A new mission to improve the monitoring of hydrology, cryosphere, oceanography, solid Earth and climate-change is therefore strongly anticipated. This evolution will provide enhanced continuity of science and services with respect to improving upon current capabilities and enabling novel science, applications and services. Achieving accurate purely satellite-based solutions on daily to weekly time scales was not possible with the previous generation of gravity missions. Indeed, at the moment this is only achievable in combination with models (Croteau et al., 2020). The Mass-change And Geosciences International Constellation (MAGIC), a collaboration between the European Space Agency (ESA) and the National Aeronautics and Space Administration (NASA) initiated over a decade ago, aims at fulfilling this new objective to improve current models and monitoring and in particular to introduce the capability to monitor and forecast extreme events like floods, droughts and other natural hazards.

The European Next-Generation Gravity Mission (NGGM), part of MAGIC, is currently in its Phase A Extension as first Mission of Opportunity in the ESA's FutureEO Programme. In the frame of the international cooperation, ESA and NASA have coordinated studies on gravity constellations to optimise the retrieval of mass-change and transport in the Earth system. The new

high spatio-temporal resolutions enable novel applications with the possibility to achieve short-term (or fast-track) gravity products in a sub-weekly basis. The collaboration aims at fulfilling the needs of international users communities, which are well expressed in the IUGG report from 2015 (Pail et al., 2015). The ESA/NASA Ad-hoc Joint Science Study Team (AJSST) contributed to the consolidation of mission requirements of the joint ESA/NASA MAGIC Mission Requirements Document (MRD) (MAGIC MRD, 2020), where recommendations from the 2015 IUGG report (Pail et al., 2015) and the 2017 US Decadal Survey (Decadal Survey, 2017) were adopted. Further recommendations based on previous work and studies (e.g., (Bender et al., 2008; Iran Pour et al., 2015; Pail et al., 2019; Purkhauser et al., 2020)), NASA/ESA Interagency Gravity Science Working Group (IGSWG) (Visser et al., 2016) and the latest advances in satellite gravimetry were also incorporated in the MAGIC MRD.

The first pair (P1) of the MAGIC Constellation will be implemented via a NASA/German Aerospace Center (DLR) fast-paced cooperation to ensure continuity of observations. The second pair (P2) will be implemented by ESA, possibly with some NASA contributions. A staggered launch approach for the two satellite pairs should provide at least 4 years of combined operations for MAGIC (MAGIC MRD, 2020). On NASA side, a Phase A study was initiated in March 2023. On ESA side, since 2020 the NGGM/MAGIC concept is investigated in two parallel industrial Phase A studies complemented by a science support study `https://www.asg.ed.tum.de/en/iapg/magic/`. In the frame of the latter, several potential architectures and mission scenarios were investigated and numerically simulated to maximise the MAGIC's scientific return. The Bender-type double-pair mission concept (Bender et al., 2008) and single/multiple pendulum configurations (Elsaka et al., 2013) were simulated in great depth. In these simulations, realistic error assumptions regarding the key payload products, in close interaction with the two parallel industry studies, were also implemented. In the frame of the science support activities for MAGIC, methodological improvements of processing strategies, optimum treatment of long-term signals, and tailored post-processing techniques were also analysed (Abrykosov et al. 2021; Abrykosov et al. 2022; Heller-Kaikov et al. 2023). The resulting simulations provided a clear overview on mission performance and scientific improvements enabled by MAGIC. Beyond the description of

simulations setup and their improvement, which is available in (Heller-Kaikov et al. 2023), this article summarizes the initial results and recommendations from the ESA science study, and focuses on the scientific applications and the improvements expected to be achieved with MAGIC.

A brief introduction on analysed constellation scenarios and adopted methodology is given in Sections 2 and 3, respectively. In Section 4, the main results and comparisons with the MAGIC MRD requirements are shown and discussed. Section 5 presents the scientific impact and applications over a set of specific thematic fields: hydrology, cryosphere, oceanography, solid Earth, climate-change and geodesy. Conclusions and recommendations for on-going and future work are finally provided in Section 6. It should be noted that, as result of technical and programmatic constraints, the current assumption for the MAGIC orbits is somewhat different than the cases presented in this paper and on-going studies address such orbit configuration. However, the results presented are fully applicable but for minor aspects that will be described in later publications.

2 CONSTELLATION SCENARIOS

In the ESA Science Support study, the analysed orbits are based on the original candidate orbits provided in (Massotti et al., 2021) and on a few additional scenarios which will be discussed hereafter. The main attention is often focused on the so-called 3d_H scenario, which is defined by a polar pair at 463 km mean altitude and 89 degrees inclination, and a second pair at 432 km mean altitude and 70 degrees inclination. The full list of analysed orbits is available in Annex A. To study the impact of different periods of near repeat orbits, or sub-cycles, and the influence of a change of altitude, several scenarios were defined. Beyond the Bender configurations, a few sun-synchronous orbit (SSO) and pendulum concepts were also introduced. For all satellite pairs, the nominal inter-satellite distance (ISD) baseline length is set to 220 km. For the scenarios 3d_H and 5d_LL additional satellite tandems were analysed, including inline tandems with a baseline length of 100 km, 150 km, and 180 km, and pendulum pairs with angles of 15, 30, and 45 degrees.

Several potential mission architectures were investigated to narrow down the trade space of the constellation, especially to provide feedback to the parallel Phase A system industry studies, and to identify an optimum set-up regarding science return, technical feasibility, and costs.

3 METHODOLOGY

The numerical closed-loop simulations performed to evaluate the compliance of various mission architectures to the given user requirements and their impact on applications are based on the full-scale gravity simulation software at the Institute of Astronomical and Physical Geodesy (IAPG), which is described in detail in (Daras et al. 2015) and (Daras 2016). The set-up of these simulations is described in detail in (Heller-Kaikov et al. 2023) as summarized briefly in what follows. As a first step, orbits of the involved satellites are computed at a sampling rate of 5 s. The orbit scenarios underlying the simulations are summarized in the Table A in the Appendix. All orbits have specific sub-cycles: after a pre-defined period of time, the satellites come close to their initial (Earth-fixed) position, with a longitudinal shift as specified in the table, thus generating a homogeneous ground track pattern within a given sub-cycle. All simulations refer to the period of time starting at January 1st, 2002. Depending on the specific analysis, monthly (31 days), weekly (7 days) or short-term (3 days) solutions are computed. Normal equations (NEQs) are assembled and solved up to a maximum spherical harmonic (SH) degree/order (d/o) of 120, 100 or 70, depending on the retrieval period.

Regarding force (background) models in the closed-loop simulation, we distinguish between models used for forward modeling ("true world"), i.e., computation of the orbits and simulation of the observations, and the gravity retrieval ("reference world"), i.e. set-up of observation equations and observation residuals. For the forward modeling the ocean tide model EOT11a (Savcenko and Bosch 2012) and the full atmosphere, ocean, hydrology, ice and solid Earth (AOHIS) signal given by the updated Earth System Model of ESA (Dobslaw et al. 2015) are used. For the computation of reference observations, the ocean tide model GOT4.7 (Ray 2008) is applied, meaning that we use the difference of the two ocean tide models EOT11a and GOT4.7 as a measure for the ocean tide background model error. In the case of the nominal processing scheme, the AO-dealiasing (AOD) product and the corresponding error estimates (Dobslaw et al. 2016) are used in the inversion, such that HIS signals are retrieved. In both forward and backward modelling the static gravity field model GOCO05s (Mayer-Gürr et al. 2015) is used, assuming that it is perfectly known. All background models extend up to d/o 120.

From the simulated orbits, synthetic High-Low (HL) and Low-Low (LL) Satellite-to-Satellite Tracking (SST) observations and residual observations are computed, using reference force models as described above. Product-noise models of the key instruments, laser ranging interferometer (LRI), accelerometers (ACC) and GNSS receiver are superimposed (their models are assumed to include the effects of all the interactions with the satellite, e.g. those with estimation and control systems for attitude and thermal stabilisation). These noise models are defined in chapter 2.1.2 of (Heller-Kaikov et al. 2023). The SH coefficients are retrieved by means of a standard least-squares parameter adjustment based on a Gauss-Markov model. The stochastic model is derived from pre-fit residuals of product-only noise simulations. Therefore, it is composed only of instrument errors and does not contain modelled temporal aliasing effects. The retrieval error $\Delta x = x - x_{HIS}$ characterizing the gravity retrieval performance of the considered simulated mission setup is computed as difference of the retrieved SH coefficients x and the SH coefficients x_{HIS} of the underlying "true" mean HIS signal of the respective period of time. In order to visualize and compare the global retrieval performance of several scenarios, we compute the degree amplitudes of their retrieval errors Δc_{nm} and Δs_{nm} in units of Equivalent Water Height (EWH, (Wahr et al. 1998)) according to

$$\sigma(n) = \frac{a\rho_e}{3\rho_w}\frac{2n+1}{1+k_n}\sqrt{\sum_{m=0}^{n}(\Delta c_{nm}^2 + \Delta s_{nm}^2)} \tag{1}$$

where a is the semi-major axis of the Earth ellipsoid, ρ_e the mean density of Earth, ρ_w the density of water, k_n are the Love numbers, and n and m represent the SH degree and order.

Detailed numerical studies have shown that the resulting performance scales with the retrieval period, respectively the number of underlying observations N according to the Gaussian error propagation rule of $\sqrt{N}$. This does not only hold for product-only error cases including system measurement errors only, but also for the full-noise cases including also temporal aliasing errors (Pail et al., 2022).

4 RESULTS AND MATCH AGAINST MAGIC REQUIREMENTS

Figure 1 shows a performance overview of the different constellation designs for a recovery period of 31 days as defined in Table A of the Appendix. These results, which are mainly based on the 3d_H scenario including realistic error models for the key instruments and tidal and non-tidal background model errors (cf. Section 3), clearly demonstrate the superior performance of Bender double-pair mission concepts over all other potential constellations, such as single-pair inline, SSO, or pendulum architectures. In case where the results include all error sources are labeled as ”full-noise”, whereas in case tidal and non-tidal background model errors are omitted they are labeled as ”product-only” solutions.

At this point it is to mention that following the approach from ESA NGGM and MAGIC MRD documents ((MAGIC MRD, 2020)), we use in this study unfiltered solutions as a performance metric. Any post-processing option would affect the mission performance in a different manner (different handling of omission and commission error, different leakage and smoothing effects, different signal dampening effects). Therefore, we consider the use of unfiltered solutions as the most unambiguous strategy. It is evident, that the difference between a single-pair and a double-pair solution will be largest for unfiltered solutions, because any type of filtering will attenuate the advantage of double-pairs regarding their increased spatial and temporal resolution and their intrinsic improved de-aliasing capabilities. Impact studies based on post-processed solution s were performed, e.g., by (Hauk and Wiese 2020) and (Wiese et al. 2022).

On top of previously mentioned benefits of flying in a Bender constellation, the altitude remains the main performance driver. In addition, in a double-pair scenario the relative contribution of the inclined pair to the total performance is more than 90% in the areas covered by both pairs (Pail et al., 2022), (Zhou et al. 2021). Therefore, a low altitude together with high-performance instrumentation of the inclined pair is absolutely essential (Pail et al., 2022). The Science Support Study in the MAGIC Phase A has also analysed different Inter-Satellite Distance (ISD) choices between satellites of the same pair. This distance represents a compromise between sensitivity (which improves with a longer distance) and spatial resolution (which degrades with greater distances). Based on performed simulations (Pail et al., 2022), ISD has an optimum for a distance

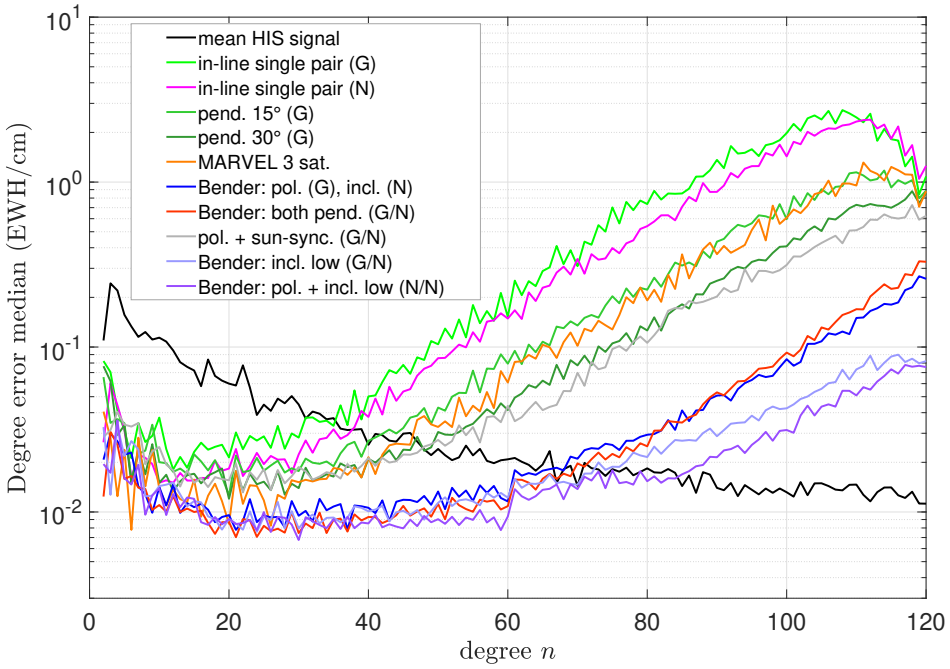

Figure 1. Degree error amplitudes of 31-day full-noise solutions from various mission constellations. ”G” means “SuperStar” (GRACE-type), N “MicroStar” (NGGM/MAGIC-type) accelerometer performance. The numbers ”15°” and ”30°” refer to the opening angle of the pendulum formation.

of 200–250 km. For this reason, it is recommended for each pair to fly with in-line formations separated by 220 km, similar as realized on GRACE and GRACE-FO.

The simulated Level-2 gravity solutions demonstrate a pronounced gain using a double-pair with respect to single pair GRACE-type scenarios, as noticeable in Figure 2, showing the improvement ratios with respect to single-pair. The greatest improvement can be seen for the ”LL” scenario (5d_LL, also labelled as ”Bender: pol. + incl. low (N/N)”). This comes as an outcome of the altitude choice being the lowest among the represented architectures. Above SH degree 40–50, for most of the alternative scenarios, improvements can go beyond 10 times higher than single-pair GRACE-type configurations or high-altitudes scenarios. In the coefficient band around SH degree 80, where the Signal-to-Noise ratio is close to one for two-pair constellations, the improvements are around 30 for Bender in-line scenario 3d_H compared to single-pair GRACE-type configurations.

The Level-2 mission performance of different architectures is compared against user requirements summarized in the MAGIC MRD (MAGIC MRD, 2020). Figure 3 depicts the cumulative

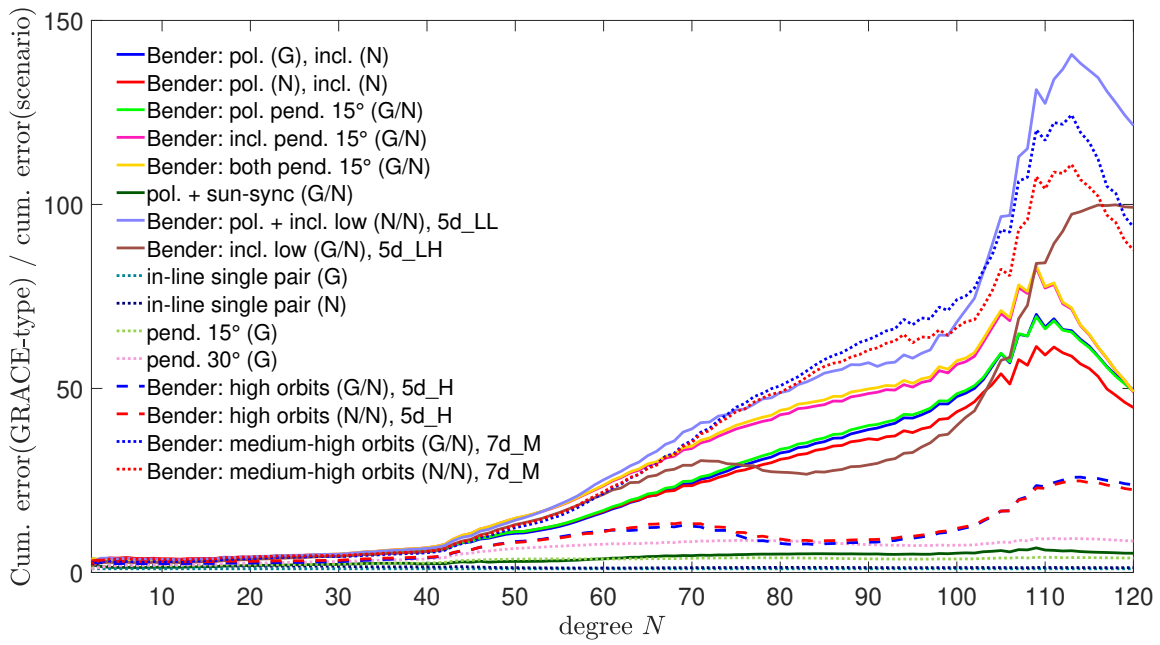

Figure 2. Ratio of the cumulative error curve of the GRACE-type single-pair and other full-noise simulated scenarios for a recovery period of 31-days. “High orbits” and ”medium-high orbits” refer to 5d_H and 7d_M scenarios, respectively. The Bender scenario with both low altitude polar and inclined pairs refer to 5d_LL, while the double pair with low inclined pair consists of 5d_LH. The “15” and “30” numbers refer to the opening angle of the pendulum formation, the letter ”G” or ”N” refers to the GRACE-like and NGGM-like accelerometer noise assumptions. More details about orbit scenarios are available in Appendix-A.

RMS curves for monthly full-noise solutions, while Figure 4 for the product-only solutions. The product-only case includes the contribution of the measurement system error, which is defined as the uncertainty of Level-2 gravity field products solely resulting from satellite instrument inaccuracies and their coupling effects at satellite (or constellation) level but excluding all other effects (e.g., tidal, and non-tidal aliasing errors). The full-noise case includes the total effect of all error sources, including tidal and non-tidal aliasing errors. In Figures 3 and 4 the MAGIC MRD threshold and target Level-2 time-varying gravity field product requirements are also plotted. It is worth to note that these requirements are adopted from the IUGG report (Pail et al., 2015), with the remark that this reports defines user needs that can be fulfilled by post-processed (e.g., filtered) solutions. The IUGG user requirements are partially based on simulations with rather optimistic assumptions on Atmosphere and Ocean (AO) background model errors, because only stochastic errors, but no deterministic (signal-related) error components were assumed. To obtain a more real-

istic assessment of the fulfilment of requirements, it would be also necessary to filter the solutions, which would reduce the cumulative errors. As defined above, this paper follows the approach from (MAGIC MRD, 2020) in which the user requirements are answered by mission performance at Level-2 which guarantees consistent traceability to the user needs being as close as possible to the geophysical signal of interest, with the possibility of further satisfying user needs via higher level post-processed products. Therefore, the IUGG user requirements comparison against the MAGIC performance can be considered as conservative. In Figures 3 and 4, and in similar comparisons in this article, post-processing was not applied in order to avoid alterations introduced by filters, which could make the comparison difficult to interpret afterwards. The mentioned points here above explain the breach of threshold requirements at low-to-medium SH degrees of the full-noise case in Figure 3. The comparison of the cumulative errors with the IUGG requirements shows that in order to meet threshold requirements and approach target requirements, a double-pair mission is required. The uncertainties of the single-pair and pendulum formations are too high to meet such requirements. The altitude turns out to have a particular influence especially for the performance at mid-to-high SH degrees. The scenarios using the 5d_H orbits (which have the highest satellite altitudes) perform poorly compared to 3d_H scenarios. The best performance is achieved by 7d_M and 5d_LL scenarios, which are characterized by low orbit altitudes. Pendulum formations do not provide a performance improvement with respect to the Bender constellations, and, moreover, they introduce a higher system complexity. Bender scenarios turned out to enable a relevant leap in performance compared to all other scenarios and to provide mission performance which satisfies the MAGIC MRD requirements except in the low-to-medium degrees.

Even if figures 3 and 4 introduce raw simulations, already looking at current results, from a pre-operational standpoint, current EO-enabled services, such as those for land, climate, ocean, and emergency management would largely benefit from improved mass-change data as available only from a constellation such as MAGIC (Massotti et al., 2022). Looking at the sub-monthly solutions and in particular at 7 days solutions (Figures 5 and 6), it is possible to find a similar behaviour with respect to the monthly solutions. In order to scale the monthly IUGG thresholds and targets to shorter retrieval periods, the requirements were scaled using a factor f_s

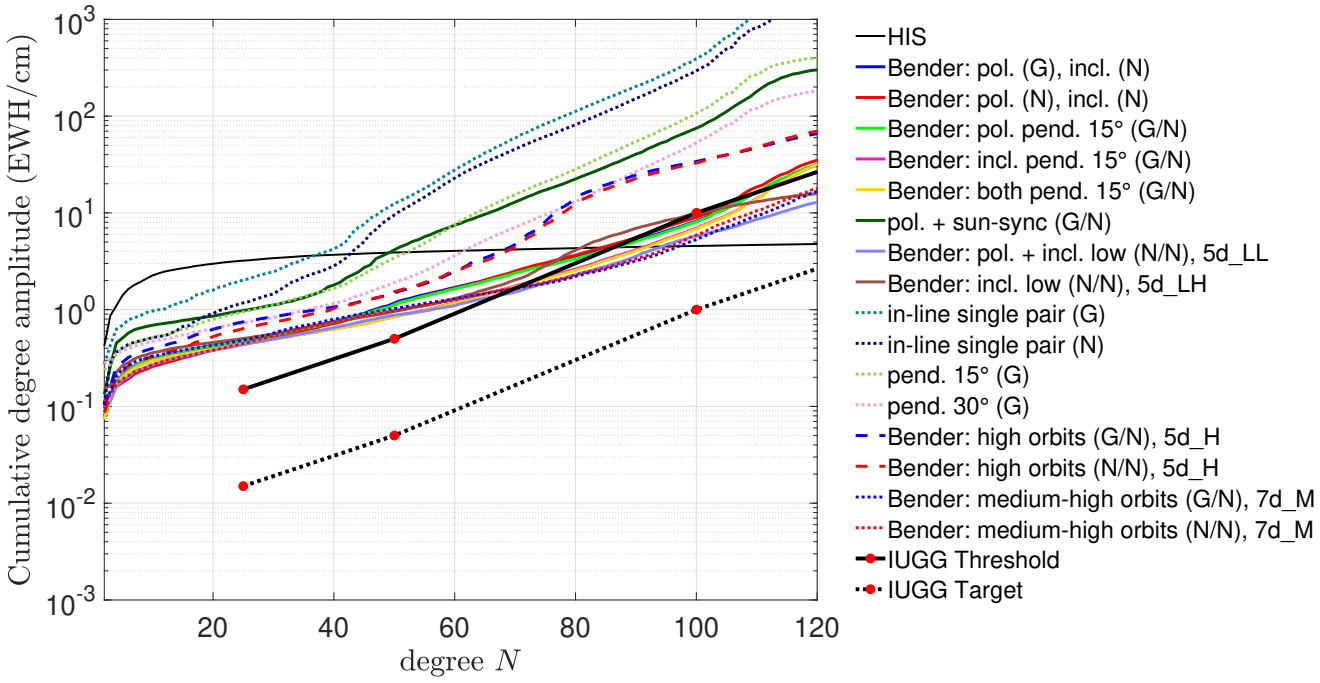

Figure 3. Cumulative RMS curves for the 31-days d/o 120 full-noise nominal simulation results, compared to the IUGG threshold and target requirements. For each individual scenario, the mean curve of the cumulative RMS curves of two subsequent 31-day solutions is shown.

$$f_s = \sqrt{\frac{31}{r_p}}, \tag{2}$$

where r_p is the retrieval period in days. In Figures 5 and 6, for each scenario, the mean curve of the cumulative RMS curves of nine subsequent 7-days solutions is provided. All simulations are generally computed up to a maximum SH d/o of 120. However, for the 7-days single pair simulations, a reduced d/o of 100 was used because of the reduced ground-track coverage which would not allow to resolve the coefficients of larger SH degrees. It is evident that GRACE-type single-pair 7-day Level-2 solutions are dominated by noise from SH degree 20 on-wards, whereas MAGIC scenarios of e.g. Bender in-line type have a Signal-to-Noise ratio of one at SH degree 70. The reduced uncertainty offered by the MAGIC constellation allows for a wide use of 7-day Level-2 time-varying gravity products even with less need of post-processing by means of filtering.

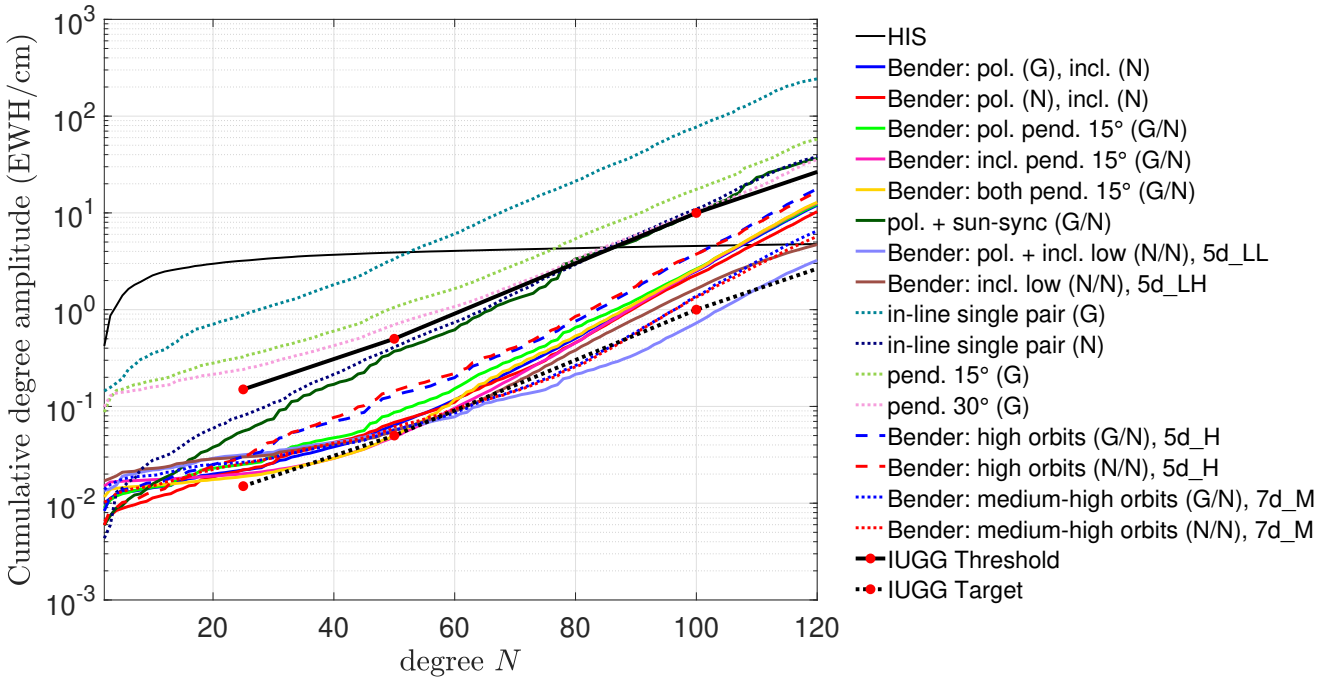

Figure 4. Cumulative RMS curves for the 31-days d/o 120 product-only simulation results, compared to the IUGG threshold and target requirements. For each individual scenario, the mean curve of the cumulative RMS curves of two subsequent 31-day solutions is shown.

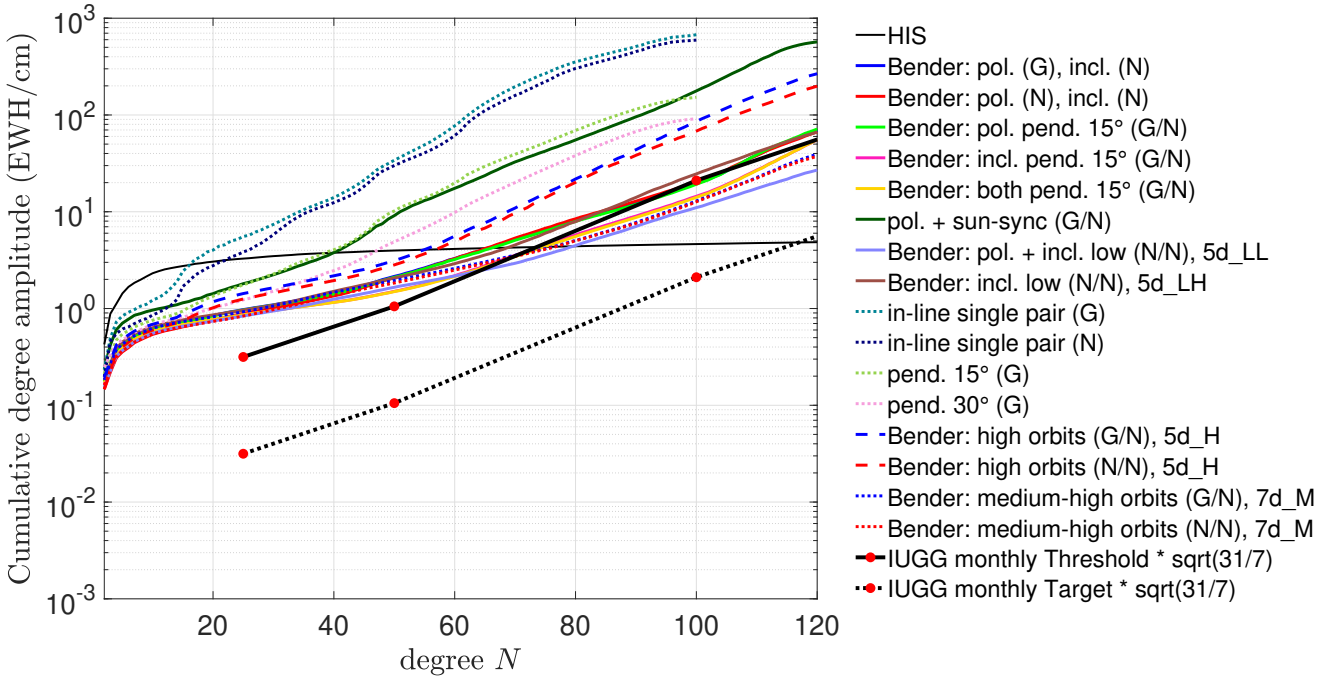

Figure 5. Cumulative RMS curves for the 7-days d/o 120 full-noise simulation results, compared to the IUGG threshold and target requirements.

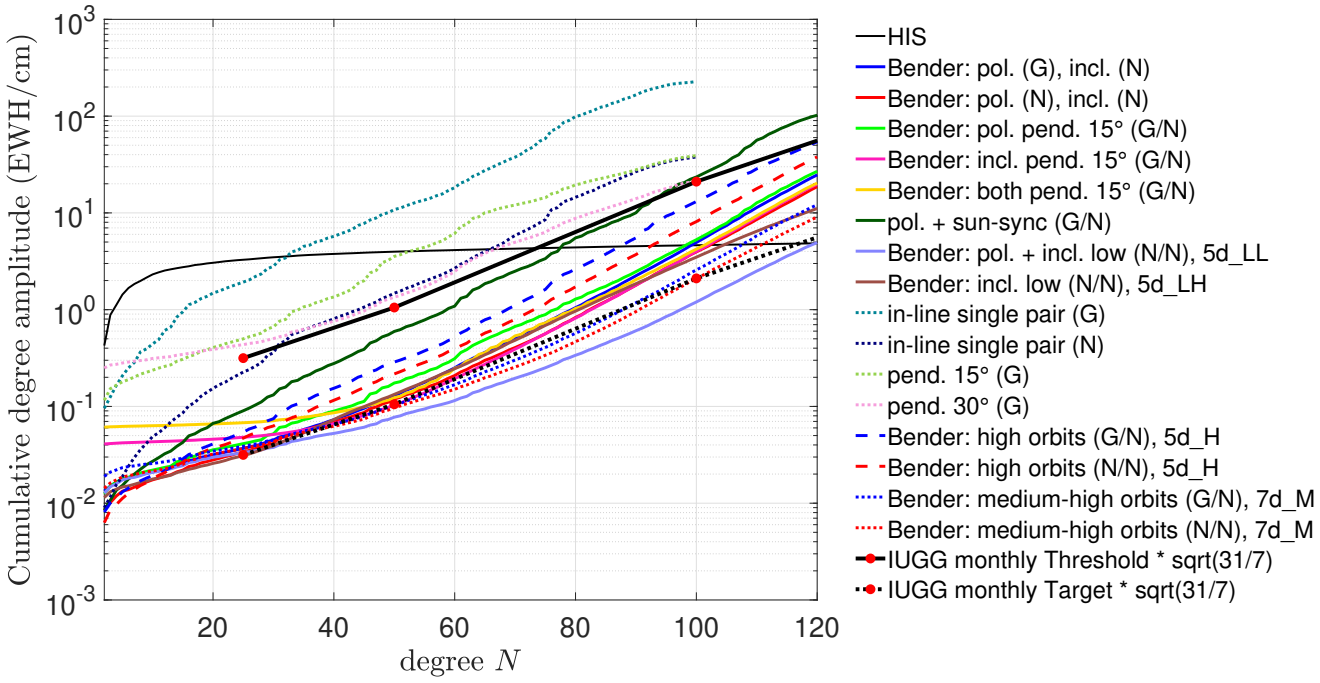

Figure 6. Cumulative RMS curves for the 7-days d/o 120 product-only simulation results, compared to the IUGG threshold and target requirements.

5 SCIENCE AND APPLICATIONS

The purpose of this section is to describe the science impact analyses of the relevant mission scenarios in different fields of applications for the mass-change data, in particular in the fields of hydrology, cryosphere, oceanography, solid Earth, climate-change and geodesy.

Sustained gravity field observations from space contribute significantly to a number of Essential Climate Variables (ECVs) as defined by the GCOS (Global Climate Observing System) programme. Among such variables, satellite gravimetry is a unique measurement technique which can retrieve global-scale data on ECVs such as ”Groundwater” and the newly adopted “Terrestrial Water Storage (TWS)” (`https://gcos.wmo.int/en/essential-climate-variables/tws`). More specifically, satellite gravimetry can provide data services for the ECV products ”Groundwater storage change” and ”TWS anomalies”.

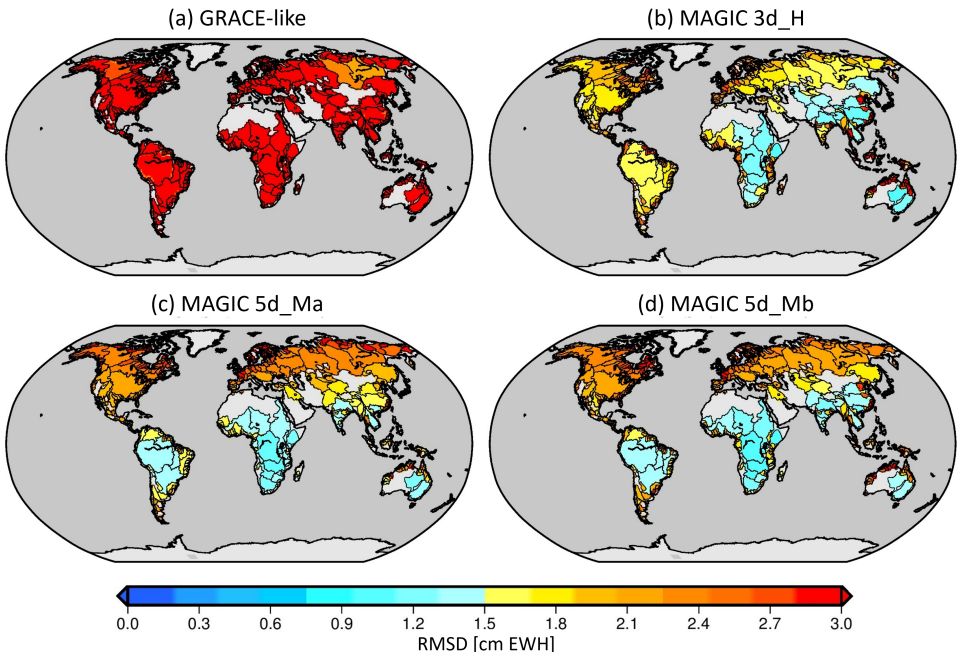

Figure 7. Temporal RMSD of basin-average water storage variations for 7-days simulation output (a: GRACE-like scenario, b: MAGIC 3d_H, c: MAGIC 5d_Ma, d: MAGIC 5d_Mb) relative to the ESM HIS reference signal, truncated at degree N=50 (i.e. 400 km spatial resolution) for 405 GRDC basins.

5.1 Hydrology

One of the most common applications of satellite gravimetry is the analysis of time series of water storage variations in hydrological units such as river basins or aquifers. These data provide fundamental information on the status of water resources, on preconditions and effects of hydrological extremes. Moreover, such data provide a valuable input for the closure of the water balance towards a comprehensive understanding of hydrological systems in response to climatic, environmental and anthropogenic changes. In spite of the unprecedented insights into hydrological dynamics that are achieved with GRACE and GRACE-FO, an even more widespread use of mass change data in water cycle studies and water resources assessments is impeded by their low resolution. For many water management applications, for instance, the value of water storage information tends to increase with its higher spatial resolution and, thus, a better match with the size of the water management units of interest, such as catchments or aquifers can be reached.

To assess the benefit of MAGIC, 7-days simulation outputs of the scenarios 3d_H, 5d_Ma and 5d_Mb (listed in Annex A), were compared to the results for a GRACE-like single pair mission.

Basin-average time series of EWH for 52 weeks were derived for 405 individual river basins defined by the Global Runoff Data Center (GRDC 2020), representing the largest river basins worldwide. The temporal root-mean-square deviation (RMSD) between the reference signal (ESA ESM HI from (Dobslaw et al. 2015)) and the simulation time series was computed for each river basin to assess the accuracy of the simulation results. Following the MRD target and threshold values for envisaged uncertainties at specific spatial resolutions (MAGIC MRD, 2020), i.e., 400 km and 260 km for the monthly time scale in the thematic field of hydrology, the spherical harmonic expansions were truncated at the degree corresponding to the desired spatial resolution, i.e. N=50 for 400 km and N=77 for 260 km. The corresponding threshold uncertainties were derived from the monthly values by error propagation following Eq. 2. This resulted in thresholds of 1.05 cm EWH for 400 km and 10.1 cm EWH for 260 km. In correspondence with chapter 4, unfiltered solutions are used to avoid the conclusions to depend on the choice of a specific filter. However, it should be noted that uncertainties of post-processed gravity field models to be later used for hydrological applications will be much smaller.

Fig. 7 shows the spatial distribution of RMSD values for all 405 river basins for the GRACE-like mission (a) and the MAGIC scenarios 3d_H (b), 5d_MA (c), and 5d_Mb (d). The improvement achieved by the MAGIC constellation is strongly visible. While the GRACE-like mission has maximum differences of more than 70 cm EWH for individual river basins and a global area-weighted mean of 10 cm EWH, the MAGIC scenarios have maximum values in the range of 4-6 cm with area-weighted means of below 2 cm. Furthermore, the different error characteristics of the orbit constellations become evident. The 5d_Ma (i.e., a lower inclination of the inclined pair compared to 3d_H) and the 5d_Mb (i.e. a lower inclination of the polar pair) scenarios appear favorable for applications in continental hydrology in lower to mid latitudes, as they show smaller residual on large parts of the continents, while the 3d_H scenario performs favorably in higher latitudes and polar regions.

A summary of the basin-average RMSD values is presented in the scatter plot in Fig. 8 (top), in which the RMSD for each of the 405 river basins is plotted against the basin size. Horizontal lines represent the MRD threshold uncertainty (1.05 cm EWH for 400 km resolution) and two

Table 1. Number of river basins for which the RMSD values of water storage variations are below given thresholds for two different spatial resolutions (400 km and 260 km), each for the GRACE-like and the MAGIC 3d_H scenarios, relative to the ESA ESM reference.

400 km, N=50			260 km, N=77		
Threshold [cm EWH]	# basins 3d_H	# basins GRACE	Threshold [cm EWH]	# basins 3d_H	# basins GRACE
3.5	90% (366/405)	2.5% (10/405)	10.1	91% (370/405)	2.5% (10/405)
2.5	67% (270/405)	0.5% (2/405)	6.1	67% (273/405)	0.5% (2/405)
1.05	0% (0/405)	0% (0/405)	2.5	17% (69/405)	0% (0/405)

additional thresholds (2.5 and 3.5 cm EWH). The vertical blue line represents the size of a spherical cap with 400 km diameter (about 125600 km^2) to roughly indicate the size of a river basin at this spatial resolution. It should be noted that this is only a rough approximation as river basins may largely deviate from a spherical shape. Signals of river basins below this size are difficult to isolate from the surroundings. It can be seen that the uncertainty requirement of the MRD can hardly be fulfilled by the unfiltered 7-days solutions for any of the river basins, including the largest ones. However, this is not surprising as the MRD thresholds were introduced for post-processed solutions. Nevertheless, also this scatter plot again stresses the strong improvement of MAGIC over a single-pair mission. Fig. 8 (bottom) shows the ratio of the RMSD of the latter compared to the MAGIC 3d_H scenario, with ratios larger than 20 especially for very small river basins. For basins around an extent of 400 km, i.e., degree N=50, the improvement by MAGIC is between 5 to 15 times. Only for very large river basins, in which the additional smoothing imposed by calculating the basin-average reduces most of the noise in both GRACE-like and MAGIC scenarios, the ratio gets smaller.

For higher spatial resolutions than 260 km (i.e., truncation at degree N=77), the threshold uncertainties provided in the MRD appear to be more relaxed, as they can be reached by all the MAGIC scenarios for most river basins (not shown). The only exceptions are basins with an area smaller than the one corresponding to a spherical cap of 260 km diameter, which are likely below the achievable spatial scale. A summary of the statistics for both spatial resolutions (400 km and 260 km) is provided in Table 1. The threshold accuracy currently given in the MRD for hydrologi-

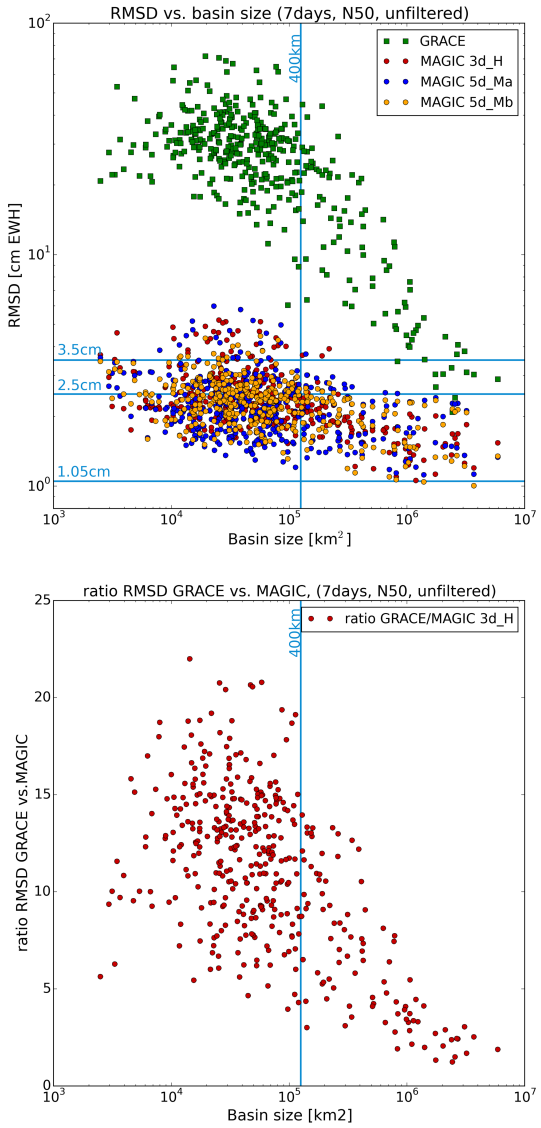

Figure 8. Top: Scatter plot of RMSD of basin averages of water storage variations for 405 GRDC river basins truncated at N=50 plotted against basin size. Blue horizontal lines indicate different uncertainty thresholds of 1.05 cm (i.e. the MRD threshold requirement), 2.5 cm and 3.5 cm EWH. The vertical blue line represents the area of a spherical cap with 400 km diameter. Bottom: Ratio of the RMSD of a GRACE-like constellation vs. the RMSD of MAGIC 3d_H constellation for each river basin.

cal applications at a comparatively high spatial resolution (10.1 cm EWH at N=77) can be fulfilled by MAGIC 3d_H (and similarly by 5d_Ma and 5d_Mb) for more than 90% of the 405 major river basins worldwide when considering unfiltered solutions. Even higher accuracies that may be required for several hydrological applications can be met in a large number of basins (Table 1, right columns). In contrast, the current MRD at the lower spatial resolution of 400 km cannot be met

Table 2. Number of basins of the Greenland (GrIS) and Antarctic (AIS) ice sheet for which the monthly threshold and target criteria are met for the four different configurations, using coefficients up to degree/order 80 and 120, corresponding to a spatial resolution of 250 km and ∼150 km, respectively.

	250 km (N=80)				∼150 km (N=120)			
Configuration	Threshold (11.6 cm EWH)		Target (1.2 cm EWH)		Threshold (105.0 cm EWH)		Target (10.5 cm EWH)	
	GrIS	AIS	GrIS	AIS	GrIS	AIS	GrIS	AIS
GRACE-like	5 /18	15 /27	0 /18	2 /27	17 /18	24 /27	5/18	14 /27
5d_Mb	17 /18	23 /27	0 /18	5 /27	18 /18	27 /27	15 /18	22 /27
5d_Ma	10 /18	19 /27	0 /18	3 /27	18 /18	24 /27	4/18	15 /27
3d_H	11 /18	21 /27	0 /18	3 /27	17 /18	27 /27	8/18	17 /27

by MAGIC for any river basin. However, relaxing this threshold to 2.5 cm or 3.5 cm EWH, which can be expected to be still acceptable for many hydrological applications, will allow for resolving TWS variations in 67% and 90% of the investigated river basins, respectively (cf. Table 1). With a GRACE-like mission this would not be possible, as even for these more relaxed numbers the RMSD for almost none of the basins stays below the thresholds. Furthermore, accuracy expectations for post-processed gravity field solutions are much higher, therefore the numbers listed above should be regarded as a relative performance improvement and not as the final uncertainties achievable with a double-pair mission for hydrological applications.

5.2 Cryosphere

The launch of GRACE in 2002 provided a breakthrough in our understanding of the glaciated regions. For the first time, mass changes of the ice sheets and glaciers could be measured directly, which revealed an imbalance of both ice sheets and all other major glacier systems (e.g., (IMBIE 2020; IMBIE 2018; Ciracì et al. 2020; Wouters et al. 2019)) and provide an invaluable data set for calibration and validation of ice sheet models (e.g., (Fettweis al. 2020; Schlegel et al. 2016)). Despite the major advances, there remains a strong demand for future improvements, to allow attribution of mass signals to individual glaciers and drainage systems, reduce signal contamination by hydrological and oceanographic mass variations, and improve the resolution and accuracy of data combination approaches (e.g. with altimetry, GNSS and other complimentary data) (Pail et

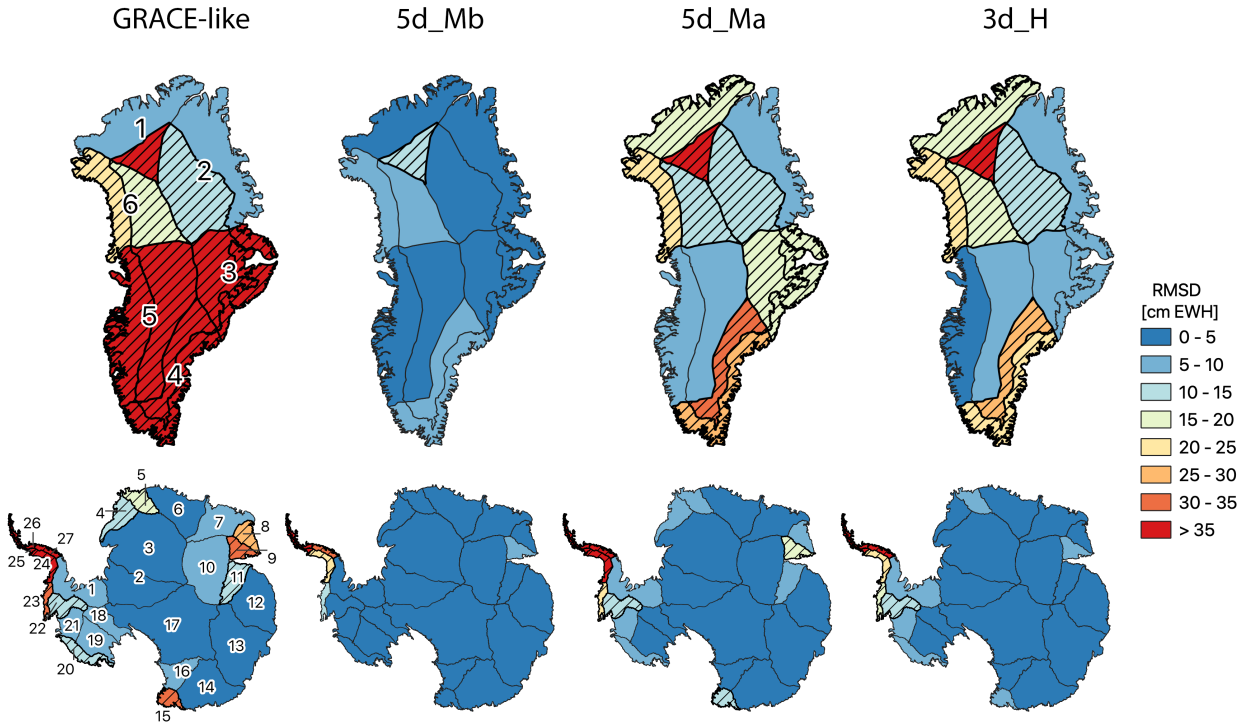

Figure 9. Temporal RMSD of 7-days simulation output for the four considered constellations and reference solution truncated at degree N=80 (i.e., 250 km spatial resolution) for Greenland (top) and Antarctica (bottom) basins. In the hatched areas, the RMSD exceeds the threshold value.

al., 2015). A future mission should therefore not only continue the current time series, especially relevant for the ice sheets where an interplay of short and multi-decadal to centennial time scales is at play, but also provide an increased spatial resolution at increased accuracy. Although the GRACE/GRACE-FO missions provide us with a clear picture of the current imbalance of the ice sheets as whole and their major sub-regions, many of the relevant processes causing this imbalance have spatial scales which remain unresolved in the current space-born gravimetric observations. On Greenland, mass loss occurs predominantly around the margins of the ice sheets. In this ablation zone, runoff of summer melt water exceeds snow accumulation, which is counteracted by a net mass gain in the interior accumulation zone. The dominant processes in these two zones are very different from a physical point of view (Noël et al. 2019). In Antarctica, the limited spatial resolution precludes us to properly separate the mass-changes on the eastern and western side of the warming Antarctic Peninsula, and of individual glacier systems in the rapidly changing Amundsen Sea Embayment.

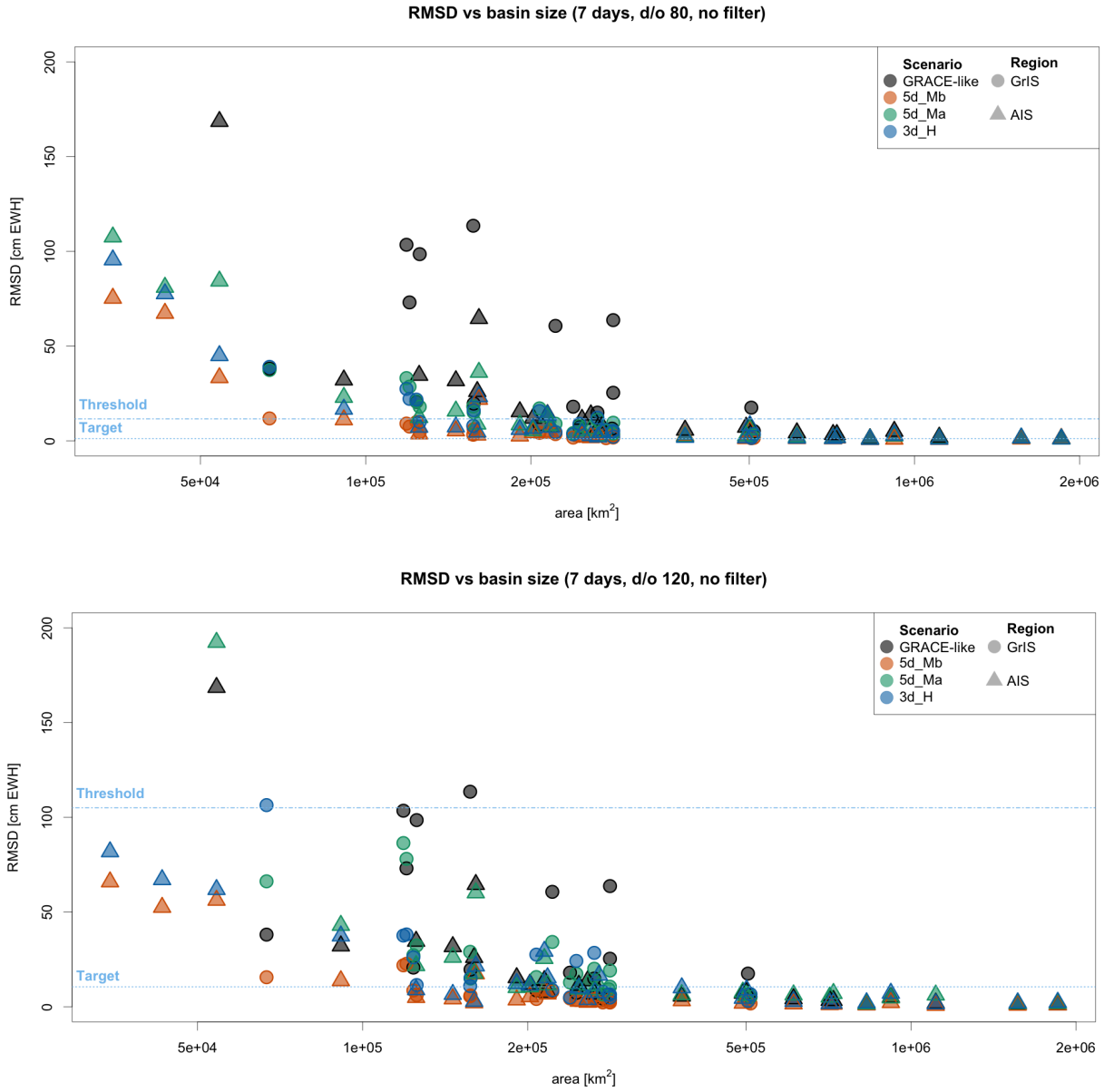

Figure 10. Top: RMS difference for each of the GrIS (dots) and AIS (triangles) basins as a function of their area, for all four configurations. Only coefficients up to degree/order 80 were used, corresponding to a spatial resolution of approximately 250 km. The dotted and dashed-dotted line indicate the (scaled) target and threshold RMS criteria at monthly time scales, respectively. Bottom: as the top figure, but for coefficients up to degree/order 120.

Here, we assess the performance in mass-change applications related to the cryosphere of different mission configurations, i.e., the one single GRACE-type pair scenario, and three Bender double-pair scenario with varying altitudes and inclinations of the polar and inclined pair (3d_H, 5d_Ma, and 5d_Mb). We focus on the ice sheets of Greenland (GrIS) and Antarctica (AIS), which

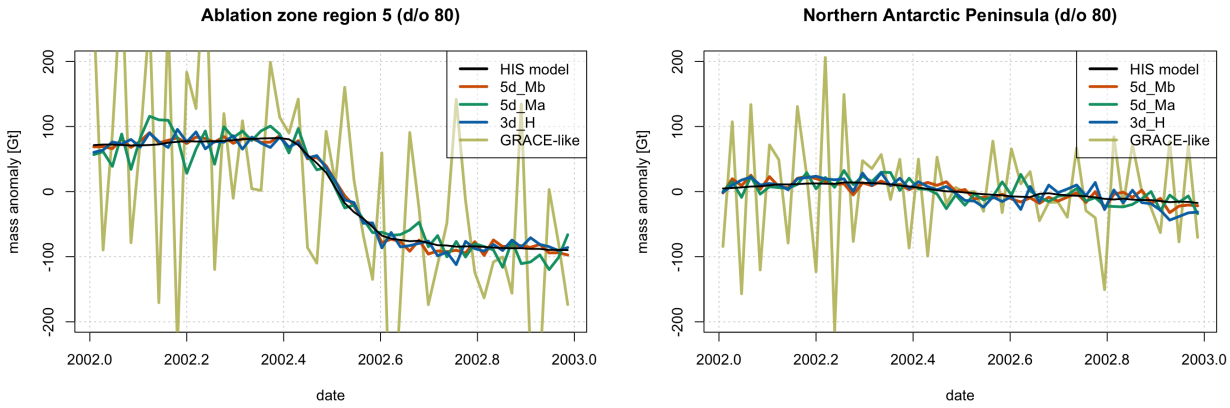

Figure 11. Mass variations in the low-elevation zone of Greenland basin 5 (left) and Northern Antarctic Peninsula (basins 24 & 25; right) simulated by the HIS model and retrieved from the 4 mission configurations simulations using the method of (Wouters et al. 2008).

were each subdivided into smaller regions, based on ice prominence, flow direction of the ice, and climatological settings. For Antarctica, the basin definition of (Zwally et al. 2012), was used, dividing the ice sheet into 27 regions. For Greenland, 6 regions were defined based on (Sasgen et al. 2012). In Greenland, these regions were further subdivided into the ablation and accumulation zone (approximated using the 2000-meter elevation contour), yielding 18 regions in total. In Antarctica, surface melt contributes minimally to mass changes, hence, a similar subdivision of the regions was not made for this ice sheet.

To retrieve the mass variations of the ice sheets and their (sub)regions, two approaches were used. First of all, the method of (Wouters et al. 2008), in which the spherical harmonics are transformed to surface mass loading anomalies. Subsequently, modelled mass anomalies in the (sub)regions are adjusted iteratively, until convergence is reached with the input mass anomalies from the simulation. Secondly, the mascon approach of (Ran et al. 2018) synthesises gravity disturbances at predefined points positioned at a specific satellite altitude, which are then converted into localized mass anomalies through a linear functional model which uses the variance-covariance matrix in its weighting. Since both methods yielded comparable results, it was decided to proceed with the method of (Wouters et al. 2008) for computational efficiency.

To assess the performance of the different mission configurations, the time series of mass variations retrieved from the simulations were compared to the truth signal, retrieved from the HIS

model from (Dobslaw et al. 2015) (i.e., without the noise component, provided up to degree/order 180). Time series for all basins were plotted for a qualitative assessment. For a quantitative analysis, the root-mean-square error was computed based on the difference between the simulated and truth time series and compared to the relevant threshold and target user requirements in the MRD (MAGIC MRD, 2020). As for the hydrology case study (section 5.1), the weekly, unfiltered spherical harmonic expansions were truncated at the appropriate degrees N and the monthly threshold (5.5 and 50 cm EWH at 250 km and 150 km, respectively) and target uncertainties (0.55 and 5 cm EWH at 250 km and 150 km, respectively) were scaled using Eq. 2. This results in threshold(/target) values of 11.6(/1.2) cm EWH at 250 km resolution ($N = 80$), and 105(/10.5) cm at 150 km resolution. For the latter, we use the maximum provided degree of $N = 120$ for the MAGIC-scenarios and $N = 100$ for the single GRACE-type pair scenario.

Figures 9 and 10, and Table 2 summarize the performance of the four configurations with respect to the threshold and target criteria for a spatial resolution of 250 km, at monthly time scales. Most notable is the improved performance of the Bender constellations with respect to a single-pair mission in the lower-latitude basins of the two ice sheets, a consequence of the addition of the inclined pair. For configuration 5d_Mb, the threshold is met for 40 out of the 45 basins. As can be seen in Figure 10, basins not passing the threshold for this configuration generally have areas smaller than approximately 62,500 (250 x 250) km^2. On Greenland, the threshold is met for all basins, except for the accumulation zone of the northernmost region 1, where the RMSD is just 1 mm above the threshold. In Antarctica, the basins exceeding the threshold are all located on the Peninsula (basins 24–27). For the 5d_Ma and 3d_H configurations, the threshold is exceeded for larger basins, at approximately 200,000 km^2. The 3d_H configuration performs slightly better than 5d_Ma, with basins 32 and 29 passing the criterion, respectively. Still, both outperform a single-pair GRACE-like configuration (20 basins), although both these configurations result in an increased RMSD in the lower-elevation zones of the northernmost basins of Greenland (1, 2 and 6), compared to the single-pair results (Fig. 9). This is a consequence of an artifact at the transition zone between the two pairs for the constellation cases caused by non-optimal gravity field recovery processing, and is a subject of future investigations. The effect is related to the fact

that in the transition zone, going from lower to higher latitudes, the data density at ±70 degrees latitude is changing abruptly from a dense ground track sampling of the constellation to a lower ground track sampling of the polar pair. Together with different noise assumptions for the polar and the inclined pair, this can cause numerical issues in the transition zone. Strategies and alternative relative weighting schemes are currently investigated to solve this issue (Pail et al., 2022). When considering the target criterion, very few basins pass, none of them located on Greenland. Again, the 5d_Mb configuration performs best, but even there only 5 out of 45 basins meet the requirements.

When considering a higher spatial resolution of ∼150 km (maximum degree/order 120), the RMSD increases for all basins for the 5d_Ma constellation, indicating that the higher coefficients carry little signal information. For the 5d_Mb and 3d_H scenarios, a reduction of up to 40% RMSD is observed in several basins (Antarctic basins 1, 3, 8, 9, 17, 24, 25, 26, and the Greenland lower-elevation basin 6). When considering the bulk basin statistics, similar conclusions hold as for the 250 km resolution. Again, the 5d_Mb configuration performs best, with all 45 basins meeting the (comparatively less stringent) threshold criterion of 50 cm RMSD. Configuration 3d_H performs almost equally well (44 basins), although it should be noted that the RMSD for most basins lies above that of 5d_Mb. The 5d_Ma and the single-pair, GRACE-like configurations show a comparable performance, in terms of number of basins passing the criterion (Figure 7 and Table 3). Compared to the 250 km results discussed earlier, more of the basins also meet the target criterion, with 5d_Mb configuration again outperforming the other constellations.

To illustrate how each of the mission scenarios may improve our understanding of the mass balance of the Greenland and Antarctic Ice Sheet, we show time series of mass anomalies in the lower-elevation zone of Greenland basin 5 and the northern part of the Antarctic Peninsula (basins 25 and 26) in figure 11. In the single-pair, GRACE-like scenario, the typical mass loss signal following summer melt is obscured by noise. In the Bender configurations, the HIS *truth* signal is accurately tracked by the simulations, with a RMSD reduction of up to 99% for the 5d_Mb scenario. Irrespective of the chosen configuration, resolving mass redistribution in basins 25 and 26 on the Northern Antarctic Peninsula remains a challenge. These elongated geographical

features have a typical width in the range of 50–100 km, well below the targeted resolution of a future mission. When retrieving mass-changes, this will cause signal leakage from one basin to the other, thereby increasing the RMSD. However, when combining the two basins, this leakage error cancels out and the RMSD drops to an order of magnitude lower than what is currently achievable with a single-pair mission.

5.3 Oceanography

Ocean bottom pressure is an unusual and uniquely valuable parameter for monitoring ocean dynamics, because it is so strongly controlled by the ocean bottom topography that it filters out many of the small scale dynamics (the ocean mesoscale) which dominate sea level variability and most other quantities (Hughes et al., 2018). This results in a mix of different behaviours: some abyssal plain regions, especially near energetically-eddying regions, are still dominated by the mesoscale. Others show coherent variability over the entire basin (enclosed by a depth contour). Above about 3,000 m depth, the ocean sidewalls show highly coherent variability (often over tens of thousands of kilometres) along depth contours, but significant structure across contours, effectively integrating the ocean circulation, and providing an efficient means to monitor the climatically important Meridional Overturning Circulation (MOC) (Hughes et al., 2018). A secondary result of this strong dynamical control is that bottom pressure signals tend to be small compared to sea level, and hence challenging to measure from space. We will express bottom pressures in cm water equivalent (cm EWH, very close to mbar, or hPa), and use the $(1/12)^\circ$ resolution NEMO ocean model described in (Hughes et al., 2018) to investigate the signals.

To illustrate the capabilities of a future mission, we choose two target signals representing a coherent basin, and a MOC-signal. The coherent basin example is taken from the Caribbean Sea (Hughes et al., 2018), a mode with dominant period of 120 days, but significant variability at other periods (standard deviation 1.56 cm EWH overall, or 1.14 cm EWH using a narrow band-pass filter).

Figure 12a shows the results of this analysis, using truncation at different spherical harmonics to illustrate the scale dependence. The analysis attempted to reconstruct 54 years of weekly

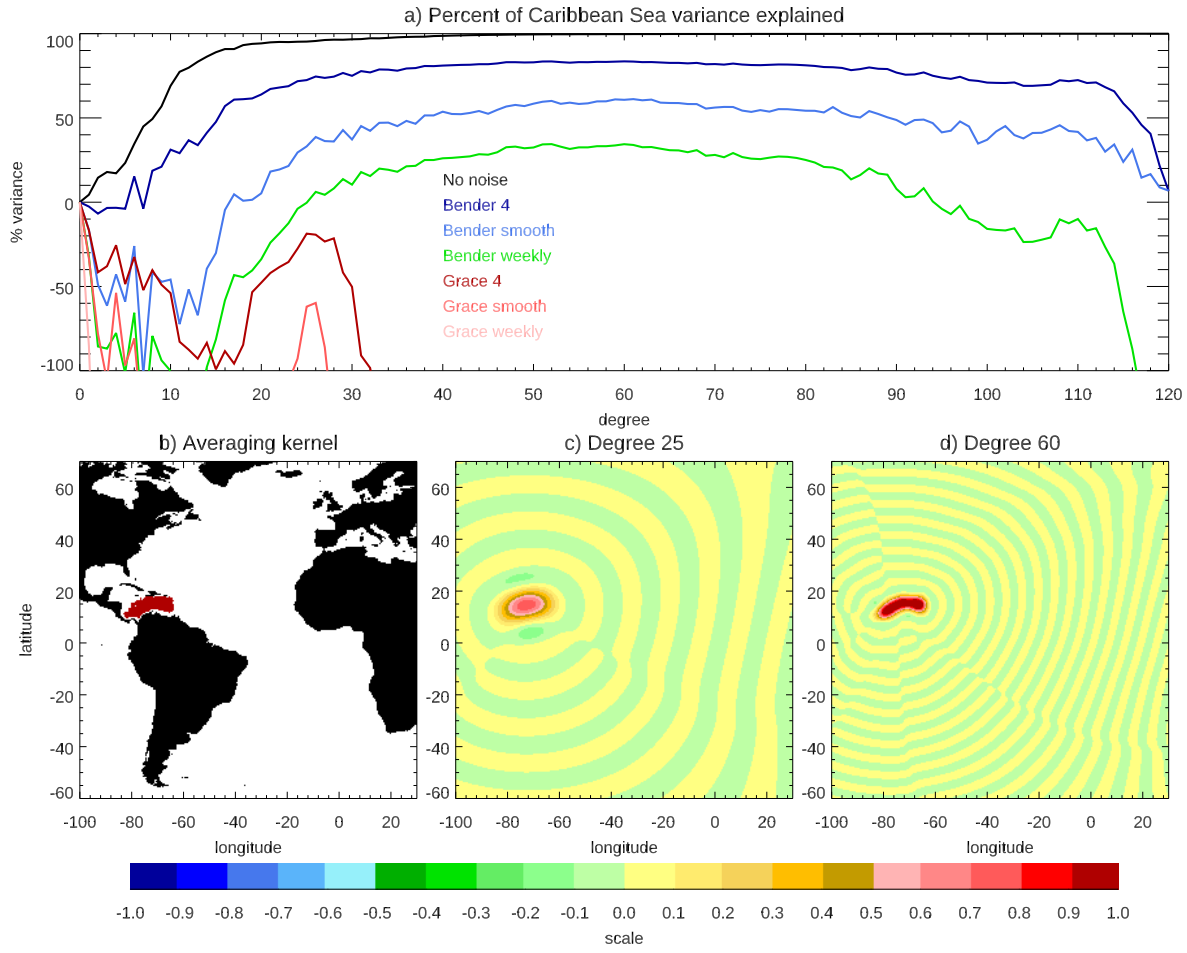

Figure 12. Retrieval of the basin-averaged bottom pressure from the Caribbean Sea, simulated under the 3d_H scenario and equivalent GRACE-like mission. Noise is taken from the weekly retrievals minus the scenario "truth", and signal from the NEMO ocean model. a) the percentage of the time series variance explained after truncating at different spherical harmonic degrees, with the noise set to zero ("no noise"), with full-noise ("weekly"), with the weekly noise divided by $\sqrt{4} = 2$ ("4"), and with 4-week averages used throughout ("smooth"). b) The averaging kernel (relative grid point weightings per unit area). c) The kernel after truncating at degree 25, and d) at degree 60.

values from NEMO, by convolving (NEMO plus noise) fields with the averaging kernel shown in Figure 12b. The noise was taken from the 3d_H scenario, and from a standard GRACE-like scenario (one year of weekly values, repeated each year). We also repeated the analysis with noise divided by $\sqrt{4} = 2$, which would approximate the effect of monthly averaging if there were no high-frequency signal; with no noise (only truncation); and with all time series reduced to 4-week means. The result of truncation in the no-noise case is, as well as adding noise from leakage, to

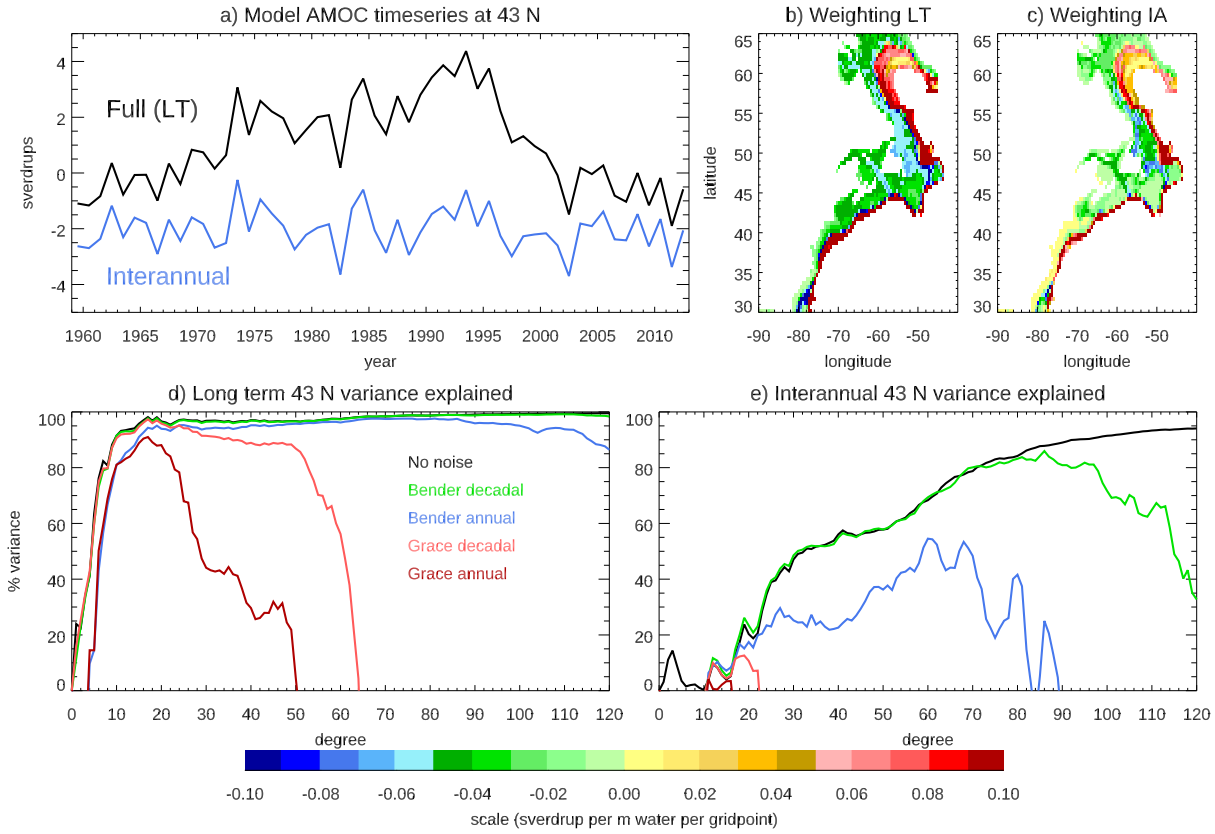

Figure 13. a) Time series of the Atlantic Overturning Circulation (overturning streamfunction at the depth of the maximum mean streamfunction) at 43° N (black), and after subtracting a running decadal mean (blue; the full time series was extended with constant values at the ends to calculate this). b) and c): weighting functions used to calculate the corresponding time series from ocean bottom pressure. d) and e): variance of the true weighted-average bottom pressure time series explained by simulated observations, after truncating at different spherical harmonic degrees, and under different observation scenarios. Weekly noise values were scaled by the square root of 52 ("annual") or 520 ("decadal") to approximate the noise at different timescales.

reduce the amplitude of the detected signal due to smoothing of the averaging kernel. To correct for this effect, a least-squares fit of the time series at each truncation on the target "true" time series was used to derive an optimal scaling factor at each truncation. This same factor (a function of harmonic degree) was applied to all other cases, thus amplifying the noise as well as the signal.

We see that the low degree harmonics do a poor job, and actually contaminate the signal in most cases. Explained variance starts to grow only after a scenario-and-noise-dependent harmonic degree is reached, and peaks at about degree 25 for the GRACE-like scenario, and about degree 60 for Bender scenarios. The corresponding averaging kernels at these degrees are shown in Fig-

ure 12c,d. There is significant leakage outside the basin at degree 25, and much better localisation at degree 60.

In this simulation, the GRACE-like scenario can never produce a useful signal, but the Bender case explains 20% of weekly variance, rising to 50% for monthly values, though still short of the 80% predicted for 4-week averages should the signal have no higher frequency variability. The Bender case is thus a dramatic improvement over the GRACE case.

In both cases, however, these are pessimistic estimates. It is known (Hughes et al., 2016) that, with carefully designed filtering, GRACE can reproduce about 40% of the signal variance for data band-passed around 120 day period. The low degree problems are probably mitigated by more careful processing design, as they would also be for MAGIC or, indeed, any other missions. However, the main conclusion of a large like-for-like improvement over a single-pair mission, and detection of a small-basin signal with amplitude around 1 cm EWH, remains clear.

A similar analysis was undertaken for the North Atlantic MOC. The MOC is measured in sverdrup (Sv) where 1 Sv $= 10^6\,\mathrm{m^3 s^{-1}}$, and Figure 13a shows a time series of 54 annual-mean values of the overturning at 43° N. A particular difficulty in this case is the potential for incidental correlations associated with large scale processes. The dynamical signal associated with the MOC is known to be related to pressures on a narrow strip of continental slope, at the ocean's western boundary, a positive MOC anomaly in the North Atlantic being associated with negative pressure anomalies in shallow water, and positive pressure anomalies between about 1200 and 3200 m depth (Elipot et al., 2014; Hughes et al., 2018). However, pressures may be correlated over a much wider region, perhaps due to a genuine dynamical connection, but perhaps also as a result of signals with common forcing, or time series with few degrees of freedom. To be sure that we are measuring the MOC itself, rather than other dynamics that are often associated with the MOC, we must take care to limit the signal to the continental slope region alone. With associated pressure signals of order 1–3 cm EWH (Hughes et al., 2018), this thin region makes a challenging target.

The model MOC time series (standard deviation 1.6 Sv) is dominated by a rise from 1959 to 1993, and a fall thereafter. This long-term signal is associated with many other global changes, including basin-average pressure changes in the Atlantic and Pacific, which means it can be re-

constructed with large-scale averages. Accordingly, we have also used an "interannual" version of the time series, in which a running decadal mean has been subtracted (blue curve in Figure 13a; standard deviation 0.72 Sv) to reduce the chance of incidental correlations.

Our target averaging kernels for the two time series are shown in Figure 13b and c. These were constructed by limiting to the region shown, and to depths shallower than 3200 m on the continental slope, fitting the MOC on the equivalent time series of bottom pressure at each grid point, and multiplying by the squared correlation to reduce the weighting for poorly correlated regions. The result is certainly not the optimal method of reconstructing the MOC, but it provides a realistic target function with the expected spatial characteristics.

As we are now using annual values, and only have noise estimates for weekly values, we divide the noise fields by $\sqrt{52}$ to simulate the noise for annual means (as noted at the end of section 3, this scaling appears to be quite effective). We also, far more speculatively, use a scaling by a factor $\sqrt{520}$ to represent decadal means. This undoubtedly omits significant systematic errors on these time scales, but is presented as an aspirational outcome for long period monitoring (decadal changes of the MOC are the most climatically interesting).

The difference between the fidelity of the two reconstructions (Figure 13d,e) is very informative. The large percentage of variance apparently explained by low degree gravity fields in the long-term case is clearly not representing a signal limited to the continental slope, but rather is the leakage of the larger-scale signal into the filtered averaging kernel.

For the interannual case, the rise in variance explained with degree, for the no noise case, is a much more plausible description of how well the kernel is realised at different resolutions. In this case, GRACE scenarios are of limited use, but the Bender case (blue) reproduces 50% of the time series variance at degree 60. While the "decadal noise" case may be too optimistic, it is worth noting that it is roughly equivalent to seeking a signal $\sqrt{10}$ times larger than the model variations (2.3 Sv rather than 0.72 Sv), and that this is a perfectly reasonable size for expected MOC changes due to climate change.

There have already been attempts at using GRACE data to measure MOC changes (Delman and Landerer 2022), with some success. It seems clear, however, that these rely on there being

auxiliary signals at larger length scales than those on the continental slope that are known to be dynamically related to the MOC (Elipot et al., 2014; Hughes et al., 2018). The higher resolution capability of a Bender configuration, providing useful extra information out to at least degree 60 and potentially about degree 85 for larger signals, is a step-change of resolution with important consequences for ocean monitoring. It leads to the possibility of genuinely isolating the continental slope signal at an accuracy which is comparable to the best in-situ measurement systems (McCarthy et al., 2020), which are themselves only representative of single latitudes, rather than of the broad latitude range that is of relevance to climate change.

5.4 Solid Earth

We have investigated to what extent a Bender configuration (3d_H) detects the coseismic gravity signal generated by a smaller earthquake compared to a GRACE-type single-pair scenario. The coseismically generated seismic waves are covered globally at magnitude M 5 or even M 4.5 by existing international seismic networks, that provide the hypocentre and fault plane, and solution of the seismic moment. Problematic is the detection of postseismic deformation as well as the intermediate phases in the earthquake cycle which give origin to a crustal deformation and a gravity signal, but do not generate seismic waves detectable by standard seismic networks (Elliott et al. 2016), (Barbot et al. 2008). These are signals for which the satellite gravity acquisitions can give complementary information to space geodetic observations, especially for faults in oceanic areas for which surface deformation cannot be detected by interferometric Synthetic Aperture Radar (InSAR) acquisitions or by Global Navigation Satellite Systems (GNSS). These movements are though slow, extending over decades or centuries, and depending on the relative velocity of the moving plates across the fault discontinuity. Therefore, a monthly or yearly time resolution of the field can be already acceptable, if a lower magnitude can be detected by MAGIC. For this reason, in the case of solid Earth applications, it is useful to consider not only the weekly spectral error curve of MAGIC in recovering the HIS signal, but also the time-averaged error curves, which are subsequently lowered with increasing integration time-interval, finally almost reaching the instrumental spectral noise curve. Given the error degree variance at spherical harmonic (SH)

degree l formulated for the solution timespan Δt_0 (original time interval), variance $\sigma^2(l, \Delta t_0)$ and Δt_1 (time interval we are scaling to), variance $\sigma^2(l, \Delta t_1)$, the scaling law is the following:

$$\sigma^2(l, \Delta t_1) = \sigma^2(l, \Delta t_0) \, \frac{\Delta t_0}{\Delta t_1} \tag{3}$$

The error curves are reduced by the approximate factor 2/100 in variance by extending the sampling from one week to one year, and, as will be shown later, lowering the curves, results in reduced minimum earthquake magnitude sensitivity for MAGIC. The earthquake is detectable if the signal spectrum in a given SH degree range is larger than the power of the degree error variance of MAGIC. Since the signal is local and not global, a spatial localization of the spectrum is required. We apply the spatio-spectral localization method in the spherical harmonic domain proposed by (Wieczorek et al., 2007), in its implementation in the SHTOOLS software (Wieczorek et al., 2018), and adopt the spherical cap as localization domain. If the localization operation is not applied, an observable earthquake spectrum results to be much lower and seemingly undetectable by the mission, which is not realistic when comparing the earthquake signal amplitude with the noise amplitude calculated on the Earth surface.

We compute the gravity field of an earthquake through the deformation in a radially layered spherical Earth model including coseismic and visco-elastic post-seismic deformation (software QSSPSTATIC (Wang et al., 2017)). The layered Earth model consists of a purely elastic lithosphere in the first 40 km, overlying a Burgers model up to 120 km, then by two Maxwell bodies separated by an interface at 660 km. Properties from 1400 km to the Earth center, are assumed constant.

The relevant parameters defining the earthquake source are the dip of the fault, the rake of slip and depth of fault. We have tested a multitude of combinations of the fault parameters, finding that the thrust fault maximizes the gravity signal. In the point-source approximation, for a given unit source-mechanism tensor, characterizing the earthquake mechanism, the tensor elements scale linearly with the seismic moment. Therefore also the body forces which constitute the source term of the fault movement scale linearly with the seismic moment. Due to the harmonicity of the gravity field, this holds true for any derived functional and/or upward continued field. Therefore, the seismic moment is the controlling parameter of the amplitude of the gravity field. This allows

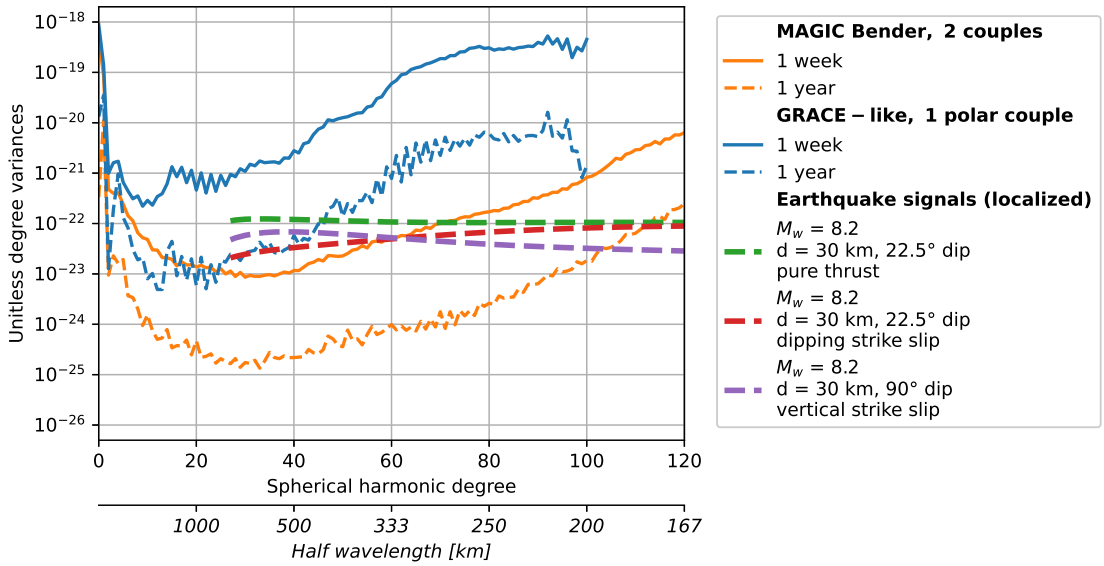

Figure 14. The localized spectrum of a M 8.2 magnitude earthquake. The MAGIC error curves are shown for 1 week and 1 year time resolution. The earthquake spectra are shown for a pure thrust, and a dipping and vertical strike-slip fault. Earthquake depth is 30 km.

to calculate the gravity field for one reference seismic moment, scaling the degree variances then for other seismic moments, without the need to recalculate the gravity field for the different fault plane solutions. In terms of degree variances of the gravity field, they scale with the square of the seismic moment. As seen in Figure 14, a magnitude M 8.2 earthquake is at the limit of observation of the single polar pair after one year of data accumulation (GRACE-like scenario), the thrust event being seen only up to degree N=45, and is well above the detection limit for the Bender double pair, up to the maximum degree N=65 with the much higher time resolution of 1 week. This is comparable with the results of (Chao and Liau 2012), who find that the earthquake signal is captured by GRACE down to M 8.3, by using an Empirical Orthogonal Function (EOF) based strategy and a step-function model over a 2-year time span: 1 year before and 1 year after the event, which can be consistently compared with the noise curves obtained with 1 year of data accumulation. We also note that (de Viron et al. 2008) provided a 30 per cent detection for events of M 8.1 and 50 per cent for M 8.8 events in GRACE gravity solutions. Their strategy used post-processed monthly solutions and an EOF based time-series analysis strategy that is expected to mitigate the effect of dealiasing residuals. A direct comparison with our approach is therefore more difficult.

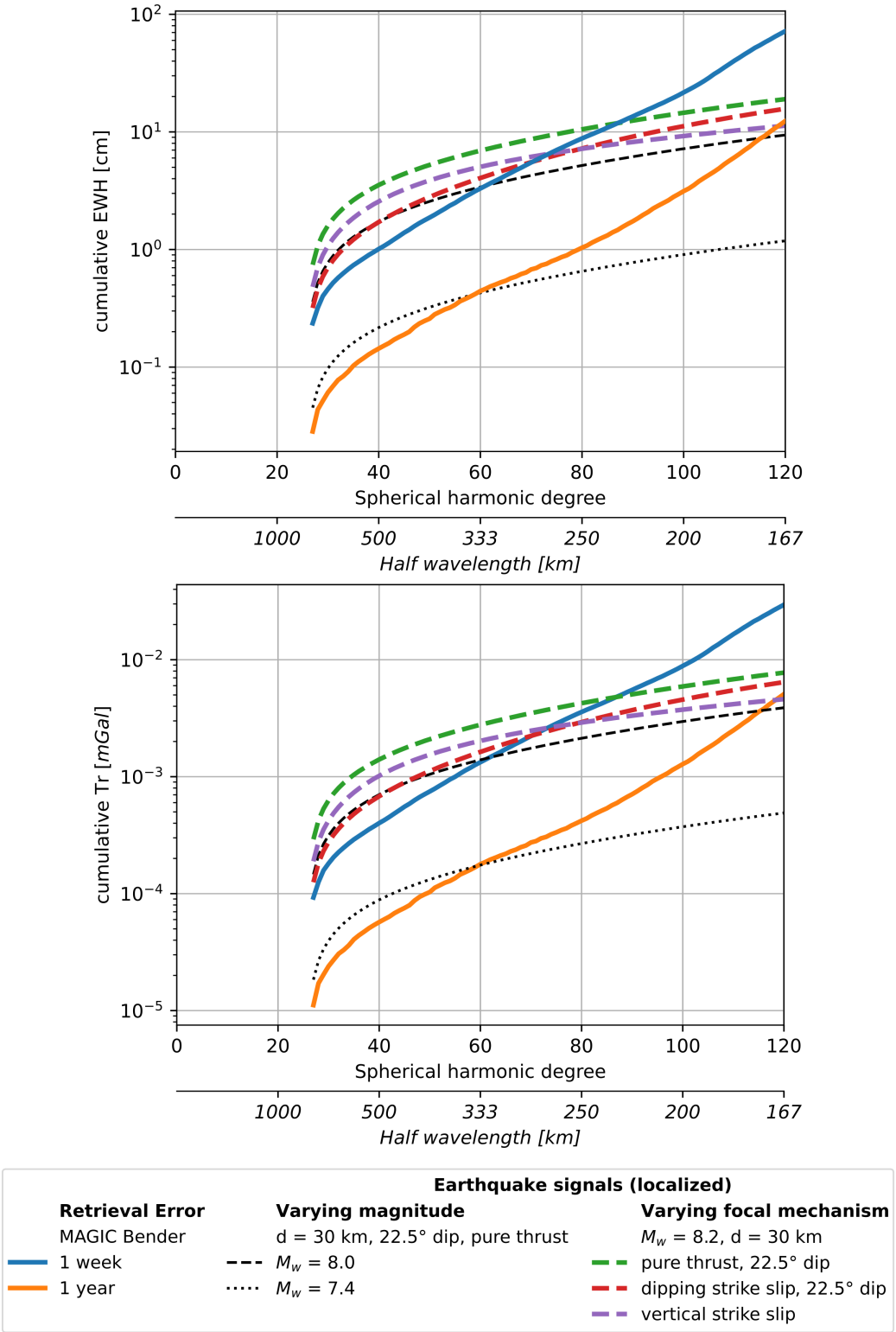

Figure 15. Spectral-domain comparison of the earthquake signals with retrieval errors in two selected scenarios (GRACE-like polar only and MAGIC Bender 3d_H). Top: EWH; Bottom: first radial derivative of the disturbing potential (T_r). The spectra in both panels are expressed in RMS per SH degree, cumulative, at nominal ground level ($r = a_{WGS84}$).

We calculate the cumulative RMS-per-degree spectra for the signal, for a number of significant simulated earthquakes, and error, in two selected scenarios (GRACE-like polar only and MAGIC Bender 3d_H). The cumulative RMS spectra at a given SH degree is representative of the signal and error RMS that can be observed in the spatial domain (i.e. by synthesizing the SH expansion on a discrete grid of computation points). We apply the aforementioned spatio-spectral localization in a 8° radius spherical cap, which results in a minimum SH degree of the localized spectra of N=27. These spectra are shown in Figure 15, expressed in two functionals derived from the SH coefficients of the potential: EWH, using the conversion defined in Eq. 1, and gravity, specifically the first radial derivative of the disturbing potential (T_r). The detection limits (the intersections of signal and error spectra) are invariant with respect to the employed functional. This serves as a control of consistent operations being carried out on the error and signal data. Signal is computed at a nominal ground level, equal by convention to the semi-major axis of WGS84. Note that the effect of a ground-tied measurement, which would be displaced by the surface displacement and be thus affected by the local gravity gradient, is removed from the modelled earthquakes signals reflecting the nature of a space-borne gravity measurement. The coseismic signal for a magnitude M 8.0 is about the limit for a weekly time resolution of MAGIC. Extending the resolution to 1 year, the lower magnitude of M 7.4 could be observed. The worldwide detection of a M 7.4 and lower earthquake is done by seismological networks as well, but the information retrieved from gravity is the combined effect of aseismic afterslip and postseismic relaxation, the modelling of which allows to further define the fault properties. In particular, the postseismic relaxation can be predicted starting from the coseismic fault dislocation and making assumptions on the crust and mantle rheology, and therefore subtracted from the observed gravity signal to determine the fault afterslip. The limit of detection for a strike slip mechanism requires a slightly greater magnitude than the M 8.0 limit for the thrust event with weekly sampling. In summary, to analyse the performance of MAGIC in detecting a gravity signal generated by an earthquake of a limiting magnitude M, we simulate the gravity signal that is generated by slip on a fault, given a seismic moment, to which a seismic magnitude is calculated. We define an earthquake to be observable if at least part of its spectral curve is above the spectral noise curve of the satellite mission. The

seismic moment could be interpreted as due to a coseismic signal, or due to a slow fault slip, and the gravity signal generated either in short time, or developing a trend distributed over time. The noise curve for a yearly time resolution has a degree variance that is smaller by a factor of 2/100 compared to a weekly time resolution, lowering the smallest observable seismic moment by the factor 0.14 (square root of 2/100), which translates into a moment magnitude reduction of 0.57 due to the difference in moment magnitude being $2/3 \cdot \log_{10}(0.14)$ (Wells and Coppersmith 1994). This is because the degree variance of the gravity signal scales with the square of the seismic moment. Comparing single and double pair configurations with weekly solutions shows that the double pair significantly lowers the detectable moment magnitude from M 8.8 to M 8.0, and increases the highest observable degree up to about 60 (333 km resolution). Given a certain earthquake magnitude, the highest observable degree is defined by the SNR being equal to one or higher. The highest resolvable degree of the earthquake depends on its magnitude: fixing the requested spatial resolution of an earthquake to 333 km at weekly sampling, which corresponds to degree 60, the Bender configuration requires the magnitude to be M 8.0, whereas the GRACE-like configuration requires the magnitude to be M 9.2. Lowering the time resolution to 1 year, the Bender configuration would detect earthquakes with magnitude M 7.4 upwards, at a spatial resolution of 333 km (degree 60). At higher degrees for this magnitude, the noise is expected to be larger than the earthquake signal. Undoubtedly, the MAGIC configuration will bring a definitive improvement compared to the present observation technology.

5.5 Climate Change

Some theories suggest that climate change might lead to an “intensification” of the global water cycle resulting in, e.g. an increase in the annual amplitude of water storage change (Huntington 2006), and/or that climate-induced changes in atmospheric circulation patterns affect the phase of annual peaks (Dunning et al. 2018). (Jensen et al. 2020) investigated the detectability of net TWS changes in the annual water cycle using satellite gravimetry. A satellite mission able to observe and quantify these changes would be beneficial in two ways: (i) Satellite gravity could be used as tool to proof (or falsify) the postulate of an intensification of the water cycle in different regions

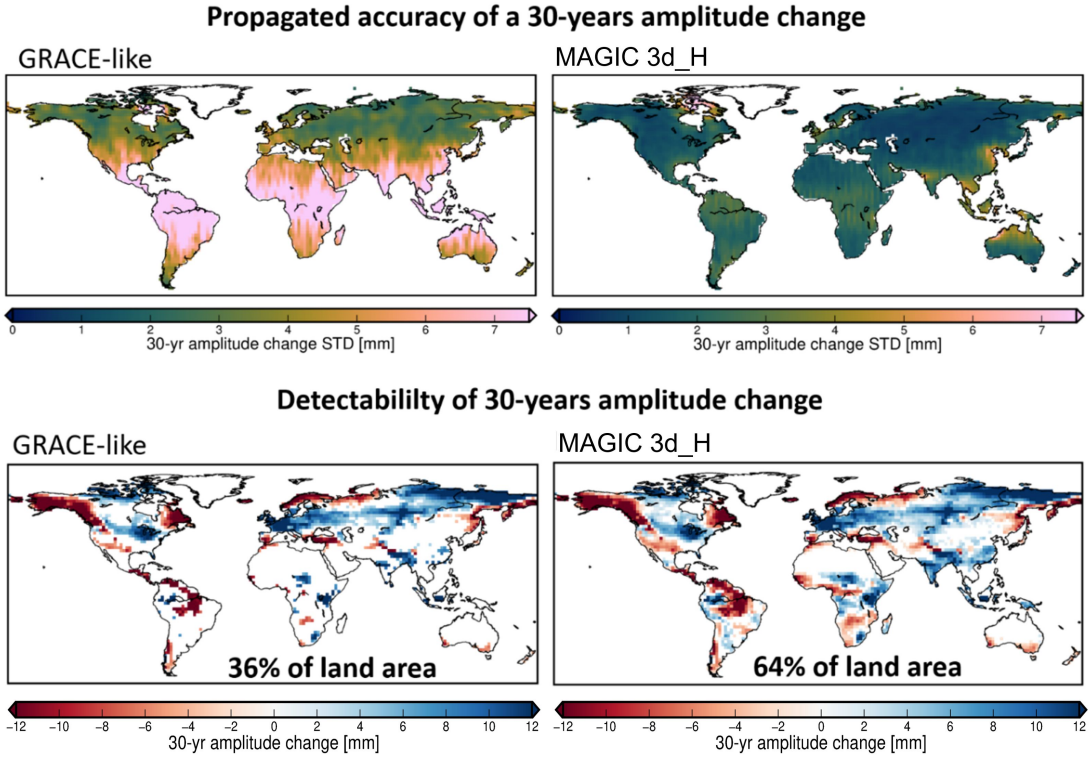

Figure 16. Top: Standard deviation of GRACE-like (left) and MAGIC (right) TWS amplitude change of annual cycle over 30 years. Bottom: Detectability of amplitude change: coloured pixels denote where projected amplitude change exceeds the magnitude of the accuracy.

of the world, and (2) the data could serve to validate whether climate models correctly simulate such changes. (Jensen et al. 2020) computed projected changes in amplitude (Fig. 7a of (Jensen et al. 2020)) and phase (Fig. 8a of (Jensen et al. 2020)) derived from an ensemble of global climate models taking part in the CMIP6 model inter-comparison project (Eyring et al. 2016).

To analyze the detectability of these projected changes of the annual cycle by current or future satellite gravimetry missions, they were compared to the achievable accuracies of these quantities from the GRACE-like and the MAGIC 3d_H simulation output following the methodology used in (Jensen et al. 2020). To this end, the grid-wise RMSD values of the simulated 7-days temporal residuals were error propagated to derive standard deviations of amplitude/phase change after 30 years. For the amplitude change these standard deviations are shown in Fig. 16 (top) for the GRACE-like (left) and for the MAGIC (right) scenarios. Here, VADER filtered solutions (Horvath et al. 2018) are used applying a relatively weak filter ($\alpha = 10$), as the unfiltered simulation

output that was used in the above chapters is not suitable for the detection of these small changes. The projected amplitude changes (from (Jensen et al. 2020)) are now challenged against these uncertainties and colored pixels in Fig. 16 (bottom) denote regions where the projected amplitude change exceeds the magnitude of the uncertainty. While, according to the simulations at hand, a GRACE-like scenario can only detect the anticipated amplitude changes in 36% of the land area after 30 years of observation, MAGIC-like scenario would be able to identify such changes in 64% of the land area. Regarding changes in the annual phase, a detectability of a 30-years phase change from the single-pair scenario can be identified in 30% of the land area and a significant increase of this portion (56% of land area) for the MAGIC scenario.

6 CONCLUSIONS AND RECOMMENDATIONS

The investigations presented in this article have demonstrated the superior performance of Bender double-pair in-line mission concepts over other potential constellation architectures, such as single-pair inline, SSO, or pendulum. As for all gravity missions, the altitude remains the main performance driver. In case of MAGIC, a low altitude together with a high-performance instrumentation of the inclined pair was shown to be crucial in satisfying the user needs. The comparison of the cumulative errors with the IUGG user requirements once again confirmed that to meet threshold requirements and approach target requirements, a Bender double-pair mission is required. The reduced uncertainty offered by the MAGIC constellation allows for a wide use of 7-day Level-2 time-varying gravity products even with less need of post-processing by means of filtering.

An inter-comparison of MAGIC and GRACE-type scenarios was performed by means of Level-2 synthetic products, consistent with the approach on answering user requirements and forward-looking to a future mass-change product with less need of post-processing. The analyses of MAGIC scenarios for different application areas in the field of geosciences confirmed that the double-pair configuration will significantly enlarge the number of observable mass-change phenomena by resolving smaller spatial scales with sufficient accuracy. In this way, also hitherto

indiscernible Earth system processes can be unraveled and quantified with the MAGIC constellation.

For hydrological applications, the MAGIC constellation will provide a significant added value because the number of hydrological units such as river basins or aquifers that can be analysed for water storage variations with certain accuracy requirements will markedly increase compared to a GRACE-like mission. For unfiltered solutions, the threshold accuracy of 10.1 cm EWH given in the current MRD at high spatial resolution (260 km) can be fulfilled by the analyzed double-pair scenarios for more than 90% of the river basins worldwide, whereas at the lower spatial resolution of 400 km, the MRD threshold accuracy may need to be relaxed to 2.5 or 3.5 cm EWH to resolve TWS variations in 67% and 90% of the investigated river basins, respectively.

The proposed MAGIC double-pair configurations will significantly improve our ability to monitor mass displacements on the ice sheets compared to what is currently possible. For example, our results show that it should become feasible to separate mass-change signals in the interior of Greenland from those in the coastal zones, and resolve small-scale mass variations in challenging regions such as the Antarctic Peninsula. The 5d_Mb configuration shows the best performance for the cryosphere applications evaluated here, with the largest number of regions of Greenland and Antarctica passing the threshold and target criteria of the current MRD.

For oceanography, the obtained results also confirmed that the MAGIC double-pair configurations produce a great improvement in ocean bottom pressure determination over a single pair GRACE-like configuration. By extending to spherical harmonic degrees that permit clear physical interpretation up to between degree 50 and about 80, depending on the signal, there is the potential to monitor meridional overturning circulation changes on time scales of years and decades. For the Caribbean Sea example analyzed here, it is shown that hitherto barely detectable signals of about 1 cm EWH RMS variability become detectable and with optimization of methods, the MAGIC configuration is expected to be able to explain 80% to 90% of its variance.

Comparing weekly solutions for single and double pair configurations, it was shown that MAGIC will significantly lower the detectable earthquake moment magnitude from M 8.8 to M 8.0, and increase the highest observable degree up to about 60 (333 km resolution). Lower-

ing the time resolution to 1 year, the Bender constellation would be able to detect earthquakes with magnitude M 7.4 upwards, at the same spatial resolution of 333 km.

Under the climate change thematic field, a GRACE-like mission can only detect the anticipated amplitude changes in 36% of the land area after 30 years of observation. MAGIC turned out be able to identify such changes in 64% of the land area. For changes in the annual phase, a detectability of a 30-years phase change from the single-pair scenario can be again identified in only 30% of the land area, while MAGIC enables a detectability over 56% of land area.

These promising results in various applications fields demonstrate, that MAGIC will have great potential to lift mass transport monitoring from space to a next level. As already mentioned in section 1, one of the main goals of MAGIC will be to provide also short-term (fast-track) products with short latency for operational service applications, such as drought and flood monitoring and prediction, and water management. The expected impact of MAGIC in this domain is currently being investigated and quantified, and related impact studies will be part of future work.

ACKNOWLEDGMENTS

This main work presented in this paper was performed in the framework of the project "NGGM/MAGIC - SCIENCE SUPPORT STUDY DURING PHASE A", ESA-ESTEC, Contract 4000134613/21/NL/FF/ab funded by the European Space Agency. The authors would like to thank the members of the consortium study for their contribution.

DATA AVAILABILITY

The Level-2 gravity field simulated data used in this publication are freely available on the International Centre for Global Earth Models (ICGEM) website: `http://icgem.gfz-potsdam.de/sl/simulated`, and shall be cited as (Daras et al. 2023). The reports and simulations' results of the acknowledged project are freely available on the following website: `https://www.asg.ed.tum.de/en/iapg/magic/`.

REFERENCES

Abrykosov, P., Sulzbach, R., Pail, R., Dobslaw, H., Thomas, M., 2021. Treatment of ocean tide background model errors in the context of GRACE/GRACE-FO data processing. *Geophys. J. Int.*, **228**(3), 1850–1865, doi: https://doi.org/10.1093/gji/ggab421.

Abrykosov, P., Murböck, M., Hauk, M., Pail, R., Flechtner, F., 2022. Data-driven multi-step self-de-aliasing approach for GRACE and GRACE-FO data processing. *Geophys. J. Int.*, **232**(2), 1006-1030, doi: https://doi.org/10.1093/gji/ggac340.

Barbot, S., Hamiel, Y. and Fialko, Y., 2008. Space geodetic investigation of the coseismic and postseismic deformation due to the 2003 Mw 7.2 Altai earthquake: Implications for the local lithospheric rheology., *J. Geophys. Res.* 113, B03403 (2008). https://doi.org/10.1029/2007JB005063

Bender, P.; Wiese, D.; Nerem, R., *A possible dual-grace mission with 90 degree and 63 degree inclination orbits*. In Proceedings of the 3rd International Symposium on Formation Flying, Missions and Technologies, Noordwijk, The Netherlands, 23–25 April 2008; ESA/ESTEC: Noordwijk, The Netherlands; pp. 1–6.

Cazenave, A., Hamlington, B., Horwath, M., Barletta, V. R., Benveniste, J., Chambers, D., Döll, P. Hogg, A.nna E., Legeais, J. F., Merrifield, M., Meyssignac, B., Mitchum, G., Nerem, S., Pail, R., Palanisamy, H., Paul, F. von Schuckmann, K., Thompson, P., 2019. Observational Requirements for Long-Term Monitoring of the Global Mean Sea Level and Its Components Over the Altimetry Era. *Frontiers in Marine Science*, **6**, 582. doi: https://doi.org/10.3389/fmars.2019.00582.

Ciracì, E., Velicogna, I. and Swenson, S., 2020. Continuity of the mass loss of the world's glaciers and ice caps from the GRACE and GRACE Follow-On missions. *Geophys. Res. Lett.*, **47**, e2019GL086926. doi: https://doi.org/10.1029/2019GL086926.

Croteau, M. J., R. S. Nerem, B. D. Loomis, and T. J. Sabaka, *Development of a Daily GRACE Mascon Solution for Terrestrial Water Storage*. Journal of Geophysical Research: Solid Earth, 2020, 125 (3), https://doi.org/10.1029/2019jb018468.

Chao, B. F. and Liau, J. R., 2019. Gravity changes due to large earthquakes detected in GRACE satellite data via empirical orthogonal function analysis. *Journal of Geophysical Research: SolidEarth*, **124**, 3024–3035. doi:https://doi.org/10.1029/2018JB016862.

Daras, I., Pail, R., Murböck, M. & Yi, W., 2015. Gravity field processing with enhanced numerical precision for LL-SST missions, *J. Geod.*, **89**(2), 99–110. doi: 10.1007/s00190-014-0764-2.

Daras, I., 2016. Gravity Field Processing Towards Future LL-SST Satellite Missions, *Deutsche Geodätische Kommission der Bayerischen Akademie der Wissenschaften*, Reihe C, Dissertationen, Heft 770, pp. 23–39. ISBN: 978-3-7696-5182-9.

Decadal Survey, 2017. *Thriving on our changing planet - Decadal Survey for Earth Science and Applications from Space*, https://doi.org/10.17226/24938, The National Academies of Sciences, Engi-

neering, Medicine, 2017.

de Viron, O., Panet, I., Mikhailov, V., Van Camp, M. and Diament, M. 2008. Retrieving earthquake signature in grace gravity solutions. *Geophysical Journal International*, **174(1)**, 14–20. doi:https://doi.org/10.1111/j.1365-246X.2008.03807.x.

Daras, I., March, G., Pail, R., Hughes, C., Braitenberg, C., Güntner, A., Eicker, A., Wouters, B., Heller-Kaikov, B., Pivetta, T. and Pastorruti, A., 2023. Level-2a simulated gravity field solutions of ESA's science support study to Mass change And Geosciences International Constellation (MAGIC) Phase A., **V. 1.0**, *GFZ Data Services*, doi:https://doi.org/10.5880/icgem.2023.005.

Delman, A. and F. Landerer, 2022. Downscaling Satellite-Based Estimates of Ocean Bottom Pressure for Tracking Deep Ocean Mass Transport, *Remote Sens.*, **14**, 1764, doi: https://doi.org/10.3390/rs14071764.

Dobslaw, H., Bergmann-Wolf, I., Dill, R., Forootan, E., Klemann, V., Kusche, J. & Sasgen, I., 2015. The updated ESA Earth System Model for future gravity mission simulation studies, *J. Geod.*, **89**(5), 505–513. doi: https://doi.org/10.1007/s00190-014-0787-8.

Dobslaw, H., Bergmann-Wolf, I., Forootan, E., Dahle, C., Mayer-Gürr, T., Kusche, J. & Flechtner, F., 2016. Modeling of present-day atmosphere and ocean non-tidal de-aliasing errors for future gravity mission simulations, *J. Geod.*, **90**(5), 423–436. doi: 10.1007/s00190-015-0884-3.

Drinkwater, M.R., Floberghagen, R., Haagmans, R., Muzi, D., Popescu, A., 2003. GOCE:ESA's first Earth Explorer Core mission. In Beutler, G., Drinkwater, M.R., Rummel, R., von Steiger, R. (eds.), Earth Gravity Field from Space – From Sensors to Earth Sciences, Space Sciences Series of ISSI, vol. 17. Kluwer Academic Publishers, Dordrecht, The Netherlands, S. 419–432, ISBN: 1-4020-1408-2.

Dunning, C.M., Black, E., and Allan, R.P., 2018. Later Wet Seasons with More Intense Rainfall over Africa under Future Climate Change. *J. Clim.* **31**, 9719–9738, doi: https://doi.org/10.1175/JCLI-D-18-0102.1.

Elliott, J., Walters, R. and Wright, T., 2016. The role of space-based observation in understanding and responding to active tectonics and earthquakes. *Nat Commun* **7**, 13844 (2016), doi: https://doi.org/10.1038/ncomms13844.

Elsaka, B., Raimondo, JC., Brieden, P. et al., *Comparing seven candidate mission configurations for temporal gravity field retrieval through full-scale numerical simulation*. J Geod 88, 31–43 (2014), https://doi.org/10.1007/s00190-013-0665-9.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E., 2016. Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, **9**(5), 1937-1958. doi: https://doi.org/10.5194/gmd-9-1937-2016.

Huntington, Thomas G., 2006. Evidence for intensification of the global water cycle: Review and synthesis,

Journal of Hydrology, **319**(1), 83–95, doi: https://doi.org/10.1016/j.jhydrol.2005.07.003.

Landerer, Felix W. and Flechtner, Frank M. and Save, Himanshu and Webb, Frank H. and Bandikova, Tamara and Bertiger, William I. and Bettadpur, Srinivas V. and Byun, Sung Hun and Dahle, Christoph and Dobslaw, Henryk and Fahnestock, Eugene and Harvey, Nate and Kang, Zhigui and Kruizinga, Gerhard L. H. and Loomis, Bryant D. and McCullough, Christopher and Murböck, Michael and Nagel, Peter and Paik, Meegyeong and Pie, Nadege and Poole, Steve and Strekalov, Dmitry and Tamisiea, Mark E. and Wang, Furun and Watkins, Michael M. and Wen, Hui-Ying and Wiese, David N. and Yuan, Dah-Ning, 2020. Extending the Global Mass Change Data Record: GRACE Follow-On Instrument and Science Data Performance. *Geophysical Research Letters*, **47**(12), doi: https://doi.org/10.1029/2020GL088306.

Elipot, S., E. Frajka-Williams, C. W. Hughes and J. Willis, 2014. The observed North Atlantic Meridional Overturning Circulation: Its meridional coherence and ocean bottom pressure. *J. Phys. Oceanogr.* **44**(2), 517-537. doi: https://doi.org/10.1175/JPO-D-13-026.1.

Fettweis, X., Hofer, S., Krebs-Kanzow, U., Amory, C., Aoki, T., Berends, C. J., Born, A., Box, J. E., Delhasse, A., Fujita, K., Gierz, P., Goelzer, H., Hanna, E., Hashimoto, A., Huybrechts, P., Kapsch, M.-L., King, M. D., Kittel, C., Lang, C., Langen, P. L., Lenaerts, J. T. M., Liston, G. E., Lohmann, G., Mernild, S. H., Mikolajewicz, U., Modali, K., Mottram, R. H., Niwano, M., Noël, B., Ryan, J. C., Smith, A., Streffing, J., Tedesco, M., Jan Van De Berg, W., Van Den Broeke, M., Van De Wal, R. S. W., Van Kampenhout, L., Wilton, D., Wouters, B., Ziemen, F., and Zolles, T., 2020. GrSMBMIP: Intercomparison of the modelled 1980-2012 surface mass balance over the Greenland Ice Sheet, *Cryosphere*, **14**, 3935–3958, doi: https://doi.org/10.5194/tc-14-3935-2020.

GRDC, 2020. Major River Basins of the World / Global Runoff Data Centre, GRDC. 2nd, rev. ext. ed. *Koblenz, Germany: Federal Institute of Hydrology (BfG).*

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

Hauk, M., Wiese, D.N., 2020. New methods for linking science objectives to remote sensing observations: a concept study using single and dual-pair satellite gravimetry architectures. *Earth and Space Science*, **7**(3), https://doi.org/10.1029/2019EA000922.

Heller-Kaikov, B., Pail, R., Daras, I. 2023. Mission Design and Processing Aspects for the Mass Change and Geoscience International Constellation (MAGIC). Submitted to *Geophys. J. Int.*, Under Review.

Horvath, A., Murböck, M., Pail, R., and Horwath, M., 2018. Decorrelation of GRACE time variable gravity field solutions using full covariance information. *Geosciences*, **8**(9), 323.

Hughes, C. W., Joanne Williams, A. Hibbert, C. Boening and J. Oram, 2016. A Rossby whistle: A resonant

basin mode observed in the Caribbean Sea, *Geophys. Res. Lett.*, **43**, 7036–7043, doi: https://doi.org/10.1002/2016GL069573.

Hughes, C. W., Joanne Williams, A. Blaker, A. Coward and V. Stepanov, 2018. A window on the deep ocean: The special value of ocean bottom pressure for monitoring the large-scale, deep-ocean circulation, *Prog. Oceanogr.*, **161**, 19–46, doi:https://doi.org/10.1016/j.pocean.2018.01.011.

IMBIE team, 2018. Mass balance of the Antarctic Ice Sheet from 1992 to 2017, *Nature*, **558**, 7709, 219-222, doi: https://doi.org/10.1038/s41586-018-0179-y.

IMBIE team, 2020. Mass balance of the Greenland Ice Sheet from 1992 to 2018, *Nature*, **579**, 7798, p. 223-239, doi: https://doi.org/10.1038/s41586-019-1855-2.

Iran Pour S., Sneeuw N., Daras I., Pail R., Murböck M., Gruber T., Tonetti S., Cornara S., Weigelt M., Van Dam T., Visser P., Teixeira da Encarnação J., 2015: *Assessment of satellite constellations for monitoring the variations in Earth gravity field (SC4MGV)*, Final Report, ESA/ESTEC Contract No. AO/1-7317/12/NL/AF.

International Union for Geodesy and Geophysics (IUGG) (Ed.)(2015) *IUGG Resolutions at the XXVI General Assembly: 22 June - 2 July 2015; Prague, Czech Republic,, Resolution 2: Future Satellite Gravity and Magnetic Mission Constellations*, XXVI General Assembly of the International Union of Geodesy and Geophysics (IUGG) (Prague 2015), Karlsruhe : IUGG Secretariat, KIT Karlsruhe Institute of Technology, https://gfzpublic.gfz-potsdam.de/rest/items/item_1626895_1/component/file_1626894/content.

International Union for Geodesy and Geophysics (IUGG) (Ed.)(2023) *IUGG Resolutions at the XXVIII General Assembly: 11 July - 20 July 2023; Berlin, Germany, Resolution 2: Sustained Terrestrial Water Storage (TWS) Monitoring by Dedicated Gravity Satellite Constellations*, XXVIII General Assembly of the International Union of Geodesy and Geophysics (IUGG) (Berlin 2023) https://iugg.org/wp-content/uploads/2023/09/2023_IUGG-GA-Resolutions.pdf.

Jensen, L., Eicker, A., Dobslaw, H., Pail, R., 2020. Emerging changes in Terrestrial Water Storage variability as a target for future satellite gravity missions, *Remote sensing*, **12**(23), 3898. doi: https://doi.org/10.3390/rs12233898.

Kornfeld, R. P., Arnold, B. W., Gross, M. A., Dahya, N. T. & Klipstein, W. M., 2019. GRACE-FO: the gravity recovery and climate experiment follow-on mission, *J. Spacecraft Rockets*, **56**(3), 931-951. doi: 10.2514/1.a34326.

Massotti L., Siemes C., March G., Haagmans R., Silvestrin P., 2021. *Next Generation Gravity Mission Elements of the Mass Change and Geoscience International Constellation: From Orbit Selection to Instrument and Mission Design*, https://doi.org/10.3390/rs13193935, Remote Sensing 2021, 13, 3935.

Massotti L., Bulit A., Daras I., Carnicero Dominguez B., Carraz O., Hall K., Heliére A., March G., Mar-

timort P., Rodrigues G., Silvestrin P., Wallace N., *Next Generation Gravity Mission design activities within the Mass Change and Geoscience International Constellation*, SPIE Remote Sensing (RS103), Berlin, Germany, 5–8 September 2022.

Mayer-Gürr, T., Pail, R., Gruber, T., Fecher, T., Rexer, M., Schuh, W.-D., Kusche, J., Brockmann, J.-M., Rieser, D., Zehentner, N., Kvas, A., Klinger, B., Baur, O., Höck, E., Krauss, S., Jäggi, A., 2015. The combined satellite gravity field model GOCO05s, *Geophys. Res. Abstracts*, **17**, EGU2015-12364, European Geosciences Union General Assembly 2015 (Vienna, Austria). doi: 10.13140/RG.2.1.4688.6807.

McCarthy, G. D., P. J. Brown, C. N. Flagg, G. Goni, L. Houpert, C. W. Hughes, R. Hummels, M. Inall, K. Jochumsen, K. M. H. Larsen, P. Lherminier, C. S. Meinen, B. I. Moat, D. Rayner, M. Rhein, A. Roessler, C. Schmid, and D. A. Smeed, 2020. Sustainable observations of the AMOC: Methodology and technology. *Rev. Geophys.* **58**(1), e2019RG000654. doi:https://doi.org/10.1029/2019RG000654.

Noël, B., van de Berg, W. J., Lhermitte, S., and van den Broeke, M. R., 2019. Rapid ablation zone expansion amplifies north Greenland mass loss, *Sci Adv*, **5**, eaaw0123, doi: https://doi.org/10.1126/sciadv.aaw0123.

Olsson, P.-A., Kristian Breili, K., Vegard Ophaug, V. et al. Holger Steffen, Mirjam Bilker-Koivula, Emil Nielsen, Tõnis Oja, Ludger Timmen, 2017. Postglacial gravity change in Fennoscandia—three decades of repeated absolute gravity observations., *Geophys. J. Int.*, **217** (2), 1141–1156, doi: https://doi.org/10.1093/gji/ggz054.

Pail R. et al., 2015. *Observing Mass Transport to Understand Global Change and to Benefit Society: Science and User Needs – An international multi-disciplinary initiative for IUGG*, https://gfzpublic.gfz-potsdam.de/pubman/item/item_1354175, Deutsche Geodätische Kommission der Bayerischen Akademie der Wissenschaften, Heft 320, ISBN 978-3-7696-8599-2, München.

Pail et al., 2019: *Additional Constellation & Scientific Analysis of the Next Generation Gravity Mission Concept (ADDCON)*, Final Report, ESA/ESTEC Contract No. 4000118480/16/NL/FF/gp.

Pail et al., 2022: *NGGM/MAGIC – Science Support Study During Phase A*, Final Report, ESA/ESTEC Contract No. RFP/3-17035/20/NL/FF/tfd. https://www.asg.ed.tum.de/iapg/magic/documents/

Purkhauser, A.F.; Siemes, C.; Pail, R., *Consistent quantification of the impact of key mission design parameters on the performance of next-generation gravity missions*. Geophys. J. Int. 2020, 221, 1190–1210, https://doi.org/10.1093/gji/ggaa070.

Ran, J., Vizcaino, M., Ditmar, P., Van Den Broeke, M. R., Moon, T., Steger, C. R., Enderlin, E. M., Wouters, B., Noël, B., Reijmer, C. H., Klees, R., Zhong, M., Liu, L., and Fettweis, X., 2018. Seasonal mass variations show timing and magnitude of meltwater storage in the Greenland Ice Sheet, *Cryosphere*, **12**, 2981-2999, doi: https://doi.org/10.5194/tc-12-2981-2018.

Rangelova, E., Wal, W. Van Der and Sideris, M.G., 2012. How Significant is the Dynamic Component of the North American Vertical Datum?. *J. Geod. Sci.*, **2** (4), 281-289, doi: https://doi.org/10.2478/

v10156-012-0005-7.

Ray, R., 2008. GOT4.7. Extension of Ray R (1999) A global ocean tide model from Topex/Poseidon altimetry GOT99.2., *NASA Tech Memo 209478*.

Reigber, C., Balmino, G., Schwintzer, P., Biancale, R., Bode, A., Lemoine, J.-M., Koenig, R., Loyer, S., Neumayer, H., Marty, J.C., Barthelmes, F., Perossanz, F., 2002. A high quality global gravity field model from CHAMP GPS tracking data and accelerometry (EIGEN-1S). *Geophys. Res. Lett.*, **29**, 14, doi: http://dx.doi.org/10.1029/2002GL015064.

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

Sánchez, L., Ågren, J., Huang, J., Wang, Y.M., Mäkinen, J., Pail, R., Barzaghi, R., Vergos, G.S., Ahlgren, K., Liu, Q., 2021. Strategy for the realisation of the International Height Reference System (IHRS) *J. of Geodesy*, **95** (33), 1432-1394, doi: https://doi.org/10.1007/s00190-021-01481-0.

Sasgen, I., van den Broeke, M., Bamber, J. L., Rignot, E., Sørensen, L. S., Wouters, B., Martinec, Z., Velicogna, I., and Simonsen, S. B., 2012. Timing and origin of recent regional ice-mass loss in Greenland, *Earth Planet. Sci. Lett.*., **333–334**, 293–303, doi: https://doi.org/10.1016/j.epsl.2012.03.033.

Savcenko, R. & Bosch, W., 2012. EOT11a - empirical ocean tide model from multi-mission satellite altimetry, *DGFI-Report No.89*.

Schlegel, N.-J., Wiese, D. N., Larour, E. Y., Watkins, M. M., Box, J. E., Fettweis, X., and van den Broeke, M. R., 2016. Application of GRACE to the assessment of model-based estimates of monthly Greenland Ice Sheet mass balance (2003–2012), *Cryosphere*, **10**, 1965-1989, doi: https://doi.org/10.5194/tc-10-1965-2016.

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

Thomas, B., Famiglietti, J., Landerer, F., Wiese, D., Molotch, N., Argus, D., 2017. GRACE Groundwater Drought Index: Evaluation of California Central Valley groundwater drought. Remote Sensing of Environment. 198, 384-392, doi: https://doi.org/10.1016/j.rse.2017.06.026.

Visser P., Bettadpur S., Chambers D., Diament M., Gruber T., Hanna E., Rodell M., Wiese D., Labreque J., Johnson T., Phillips B.,Haagmans R.,Massotti L., Siemes C., *Towards a sustained observing system for mass transport to understand global change and to benefit society*, NASA/ESA Interagency Gravity Science Working Group (IGSWG), Doc. nr.: TUDIGSWG-2016-01.

Wahr, J., Moleneer, M. & Bryan, F., 1998. Time variability of the Earth's gravity field: hydrological and oceanic effets and their possible detection using GRACE, *J. Geophys. Res. (Solid Earth)*, **103**(B12).

30205-30229, doi:10.1029/98jb02844.

Wang, R., Heimann, S., Zhang, Y., Wang, H., Dahm, T., 2017: *Complete synthetic seismograms based on a spherical self-gravitating Earth model with an atmosphere–ocean–mantle–core structure*. Geophysical Journal International 210, 1739–1764, doi: https://doi.org/10.1093/gji/ggx259.

Wells, D. L. and K. J. Coppersmith, 1994. New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement, *Bull. Seismol. Soc. Am.*, **84**, 974–1002, doi: https://doi.org/10.1785/BSSA0840040974.

Wieczorek, M.A., Simons, F.J., 2007: *Minimum-Variance Multitaper Spectral Estimation on the Sphere*. J Fourier Anal Appl, 13, 665–692, doi: https://doi.org/10.1007/s00041-006-6904-1.

Wieczorek, M.A., Meschede, M., 2018: *SHTools: Tools for Working with Spherical Harmonics*. Geochem. Geophys. Geosyst. 2018, 19, 2574–2592, doi: https://doi.org/10.1029/2018GC007529.

Wiese, D.N., Bienstock, B., Blackwood, C., Chrone, J., Loomis, B.D., Sauber, J., Rodell, M., Baize, R., Bearden, D., Case, K., Horner, S., Luthcke, S., Reager, J.T., Srinivasan, M., Tsaoussi, L., Webb, F., Whitehurst, A., Zlotnicki, V., 2022. The Mass Change Designated Observable Study: Overview and Results. *Earth and SPACE Science*, **9** (8), doi: https://doi.org/10.1029/2022EA002311.

Wouters, B., Chambers, D., and Schrama, E. J. O., 2008. GRACE observes small-scale mass loss in Greenland. *Geophys. Res. Lett.*, **35** (20), doi: https://doi.org/10.1029/2008GL034816.

Wouters, B., Gardner, A.S., and Moholdt, G., 2019. Global Glacier Mass Loss During the GRACE Satellite Mission (2002-2016). *Frontiers in Earth Science*, **7** (96), doi: https://doi.org/10.3389/feart.2019.00096.

Zwally, H. J., Giovinetto, M.B., Beckley, M.A. and Saba, J.L., 2012. Antarctic and Greenland Drainage Systems, GSFC Cryospheric Sciences Laboratory, http://icesat4.gsfc.nasa.gov/cryo_data/ant_grn_drainage_systems.php.

Zhou, H., Luo, Z., Zhou, Z., Yang, F., Pail, R., Tu, L., Yeh, H.-C. and Yang, S., 2021. What Can We Expect from the Inclined Satellite Formation for Temporal Gravity Field Determination? In Surveys in Geophysics (Vol. 42, Issue 3, pp. 699–726). Springer Science and Business Media LLC. https://doi.org/10.1007/s10712-021-09641-9.

APPENDIX A: MAGIC ORBIT SCENARIOS

Orbits sets for inclined (IP or P2) and polar (PP or P1) pairs. The ID shows the number of sub-cycle days for which the set is optimized and an additional information about the altitudes: (M)id, (H)igh. Note that the semi-major axis is reduced by 6378 km for highlighting differences in altitude. The other columns provide information about the homogeneity of the ground track patterns; more details are described in (Massotti et al., 2021) and (MAGIC MRD, 2020).

ID	IP Alt. [km]	IP Inc. [degrees]	PP Alt. [km]	PP Inc. [degrees]	hl IP [-]	hl PP [-]	Lon. Shift IP [degrees]	Lon. Shift PP [degrees]	Sub-cycles [days]
3d_M	409	70	440	89	1.368	1.383	2.308	2.384	2, 3, 8, 11, 30
3d_H	432	70	463	89	1.451	1.449	-3.076	-3.067	3, 7, 31
5d_Ma	396	65	434	89	1.397	1.383	-1.499	-1.458	2, 3, 5, 13, 18, 31
5d_Mb	397	70	425	87	1.168	1.167	0.736	0.733	2, 5, 27, 32
5d_H	465	75	488	89	1.185	1.190	0.762	0.781	4, 5, 29
7d_M	389	70	417	87	1.238	1.253	0.743	0.786	2, 7, 30
7d_H	432	70	463	89	1.218	1.226	0.672	0.692	3, 7, 31
SSO for 3d_H	477	97	463	89	1.454	1.449	-3.097	-3.067	3, 7, 31
SSO for 7d_H	477	97	463	89	1.201	1.226	0.622	0.692	3, 7, 31
5d_LL	344	70	376	89	1.423	1.410	-1.671	-1.628	1, 2, 5, 12, 29
5d_LH	344	71.5	492	89	1.169	1.172	-0.732	-0.790	5, (32-31)
U3d5d_H	432	70	492	89	1.451 (3d)	1.172 (5d)	-3.076 (3d)	-0.790 (5d)	IP: 3, 31; PP: 5, 31
U5d_H	460	70	492	89	1.061 (5d)	1.172 (5d)	-0.284 (5d)	-0.790 (5d)	IP: 5; PP: 5, 31
U3d_H	402	65	463	89	1.382	1.449	2.380	-3.067	IP: 3, 29-30; PP: 3, 7, 31

This paper has been produced using the Blackwell Scientific Publications GJI LaTeX2e class file.