ADVANCING EARTH AND SPACE SCIENCE

Earth and Space Science

RESEARCH ARTICLE

10.1029/2022EA002311

Key Points:

- A Science and Applications Traceability Matrix to satisfy the 2017 Earth Science Decadal Survey Mass Change Science Objectives was developed
- A value framework process was used to identify and evaluate high-value mass change observing systems for implementation within this decade
- High value observing systems provide continuity with Gravity Recovery and Climate Experiment Follow-On and a pathway for improved resolution and accuracy in resolving mass change

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

D. N. Wiese,
david.n.wiese@jpl.nasa.gov

Citation:

Wiese, D. N., Bienstock, B., Blackwood, C., Chrone, J., Loomis, B. D., Sauber, J., et al. (2022). The mass change designated observable study: Overview and results. *Earth and Space Science*, *9*, e2022EA002311. https://doi.org/10.1029/2022EA002311

Received 3 MAR 2022
Accepted 7 JUN 2022

Author Contributions:

Conceptualization: D. N. Wiese, B. Bienstock, D. Bearden, F. Webb
Formal analysis: D. N. Wiese, C. Blackwood, J. Chrone, B. D. Loomis, J. Sauber, M. Rodell, K. Case
Funding acquisition: L. Tsaoussi, F. Webb
Investigation: C. Blackwood, J. Chrone, B. D. Loomis, J. Sauber, M. Rodell, S.

The Mass Change Designated Observable Study: Overview and Results

D. N. Wiese[1], B. Bienstock[1], C. Blackwood[1], J. Chrone[2], B. D. Loomis[3], J. Sauber[3], M. Rodell[3], R. Baize[2], D. Bearden[1], K. Case[1], S. Horner[4], S. Luthcke[3], J. T. Reager[1], M. Srinivasan[1], L. Tsaoussi[5], F. Webb[1], A. Whitehurst[5], and V. Zlotnicki[1]

[1]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA, [2]NASA Langley Research Center, Hampton, VA, USA, [3]NASA Goddard Space Flight Center, Greenbelt, MD, USA, [4]NASA Ames Research Center, Moffett Field, CA, USA, [5]NASA Headquarters, Washington, DC, USA

Abstract The 2017–2027 United States National Academy of Sciences Decadal Survey (DS) for Earth Science and Applications from Space identified Mass Change (MC) as one of five Designated Observables (DOs) having the highest priority in terms of Earth observations required to advance Earth system science over the next decade. In response to this designation, NASA initiated several multi-center studies, with the goal of recommending observing system architectures for each DO for implementation within this decade. This paper provides an overview of the Mass Change Designated Observable (MCDO) Study along with key findings. The study process included: (a) generation of a Science and Applications Traceability Matrix (SATM) that maps required measurement parameters to the DS Science and Applications Objectives; (b) identification of three architecture classes relevant for measuring mass change: Precise Orbit Determination (POD), Satellite-Satellite-Tracking (SST) and Gravity Gradiometry (GG), along with variants within each architecture class; and (c) creation of a Value Framework process that considers science value, cost, risk, schedule, and partnership opportunities, to identify and recommend high value observing systems for further in-depth study. The study team recommended the implementation of an SST architecture, and identified variants that simultaneously (a) satisfy the baseline measurement parameters of the SATM; (b) maximize the probability of providing overlap with the Gravity Recovery and Climate Experiment Follow-On (GRACE-FO) mission currently in operation, accelerating science return from both missions; and (c) provide a pathway towards substantial improvements in resolution and accuracy of mass change data products relative to the program of record.

Plain Language Summary This study provides an overview of the Mass Change Designated Observable Study. The goals of the study were to recommend observing systems for NASA to implement within this decade to measure Earth system mass change, after it was identified in the United States National Academy of Sciences Decadal Survey as one of the five most important observations to advance Earth system science. Mass change observations are critical to understanding changes in sea level, the health of the ice sheets and glaciers worldwide, and changes in freshwater availability across the globe. The study team recommended that NASA implement an architecture similar in nature to its two predecessor missions: the Gravity Recovery and Climate Experiment (GRACE), and GRACE Follow-On (GRACE-FO). This will maximize the chances that there will not be a data gap between GRACE-FO (currently operational) and the next observing system. Further, the study team recommended collaborating with potential international partners to add more satellites to this architecture, in order to improve resolution of Earth system mass change in space and time.

1. Introduction

In January 2018, the United States National Academy of Sciences released the 2017–2027 Decadal Survey (DS) for Earth Science and Applications from Space (National Academies, 2018). The report identified five Designated Observables (DOs) [(a) Aerosols, (b) Clouds, Convection and Precipitation, (c) Mass Change, (d) Surface Biology and Geology, (e) Surface Deformation and Change] as having the highest priority in terms of Earth observations required to advance Earth system science over the next decade. NASA responded by initiating multi-center studies to identify high value observing system architectures for near-term implementation to make the required observations. These DOs are now considered the core components of NASA's Earth System Observatory to be implemented within the current decade.

Luthcke, J. T. Reager, M. Srinivasan, V. Zlotnicki
Methodology: D. N. Wiese, B. Bienstock, C. Blackwood, J. Chrone, B. D. Loomis, J. Sauber, M. Rodell, D. Bearden, K. Case, S. Luthcke, V. Zlotnicki
Project Administration: B. Bienstock, S. Horner
Software: D. N. Wiese, B. D. Loomis
Supervision: D. N. Wiese, B. Bienstock, D. Bearden, L. Tsaoussi, F. Webb, A. Whitehurst
Validation: B. D. Loomis, R. Baize, S. Horner, S. Luthcke
Writing – original draft: D. N. Wiese, J. Chrone, B. D. Loomis, M. Rodell
Writing – review & editing: D. N. Wiese, B. Bienstock, C. Blackwood, J. Chrone, B. D. Loomis, J. Sauber, M. Rodell, R. Baize, S. Horner, S. Luthcke, J. T. Reager, M. Srinivasan, L. Tsaoussi, F. Webb, A. Whitehurst, V. Zlotnicki

The designation of MC as a DO comes against the backdrop of a near-continuous 20-year climate data record of Earth system mass change established by the pioneering Gravity Recovery and Climate Experiment (GRACE; 2002–2017) mission (Tapley et al., 2019), and the currently operating GRACE Follow-On (GRACE-FO) mission (2018-Present) (Landerer et al., 2020). From the DS, the foundational basis of the MC measurements is to "ensure continuity of measurements of groundwater and water storage mass change, land ice contributions to sea-level rise, ocean mass change, ocean heat content (when combined with altimetry), glacial isostatic adjustment, and earthquake mass movement." MC measurements are intended to extend the climate data record beyond the life of GRACE-FO, while addressing eight of the DS's Most Important Science Objectives.

The focus of this paper is to provide an overview of the MCDO study along with major results and findings. The core element of the study framework is a Science and Applications Traceability Matrix (SATM) that maps the science objectives posed in the DS to required measurement parameters for MC (Section 2). Identification and classification of relevant architectures and technologies to make the required measurements is subsequently discussed (Section 3), along with a process to map the performance of those architectures relative to the SATM through the generation of science value scores (Section 4). A value framework process (Section 5) that considers science value, cost, schedule (including likelihood of having overlap with the GRACE-FO mission), risk, and partnership opportunities is then used to identify a small subset of observing systems for further in-depth study (Section 6).

2. Science and Applications Traceability Matrix

The purpose of an SATM is to establish the motivation for a mission, linking desired scientific and practical objectives to recommended measurement parameters that will drive mission design and data system decisions. A SATM can be further used to evaluate the consequences of instrument changes and descope options. The MC SATM (Table S1) includes both "baseline" and "goal" measurement parameters, encompassing the range of guidance provided in the DS, from the minimum requirements for satisfying objectives to more aspirational desires such as closing water budgets over headwater catchments. In cases where the DS objectives were ambiguous, our expert team interpreted the document with substantial input from the relevant science and applications communities. Of particular importance, the DS emphasized continuity of the Earth system mass change data record as a key goal. As a result, we determined that the quality of measurements constituting the current program of record should define baseline measurement parameters for a future MC observing system. Further, it is clear that such a baseline observing system would contribute meaningfully to the DS objectives with which MC is aligned.

The MC SATM contains 15 science and applications objectives taken directly from the DS spanning three focus areas: Climate Variability and Change, Global Hydrological Cycle and Water Resources, and Earth Surface and Interior (Figure 1). The measurement variables that define solution quality are spatial resolution, temporal resolution, and accuracy. MC and its predecessors, GRACE and GRACE-FO, are unusual in that these three variables exist within one trade-space. That is, for a given set of satellite observations, one of these three variables can be preferentially enhanced at the expense of the other two by modifying the data processing algorithms. For each of the 15 objectives we identified one of the three as the key variable, where improvements would be most beneficial to achieving that objective. In Section 4, we discuss how those choices influenced the scoring and ranking of architectures.

An innovation of the MC SATM is the identification of a Utility Score for each objective, which describes the relative importance of MC observations to achieving a given objective, considering both the (un)availability of the required observations from alternative sources and the suitability of the MC observations to address the objective. Utility Scores (*Very Low; Low; Medium; High*) were combined with the DS-prescribed Importance (*Important; Very Important; Most Important*) to derive a weighting for each objective which was later applied in the evaluation of potential architectures (Section 4).

In addition to science, the DS also emphasized practical applications for MC, most notably related to groundwater resources and drought. A Mass Change Applications Team (MCAT) was therefore established with the charge of improving understanding of the informational needs of the applied science community and agencies and industries that could benefit from MC data products. That knowledge would then be incorporated into the SATM and a separate Community Assessment Report requested by NASA's Applied Sciences Program, toward the ultimate goal of maximizing the societal benefits of a MC mission. The MCAT began by identifying applications-related

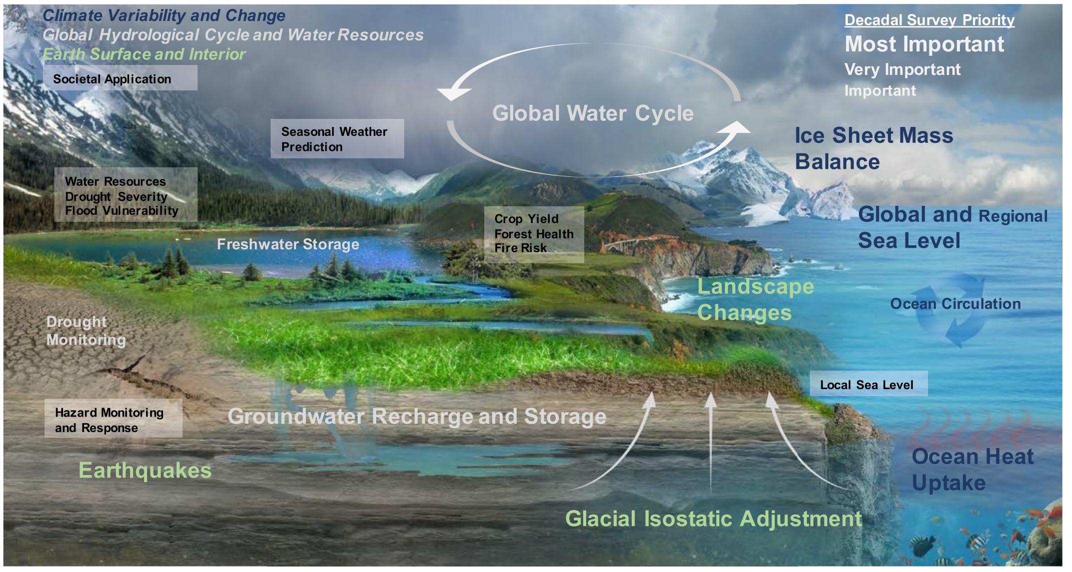

Figure 1. Science and applications objectives for Mass c=Change from the Decadal Survey.

goals in the DS and current practical uses of data products derived from GRACE and GRACE-FO. These include water resources assessment (e.g., Famiglietti et al., 2011; Richey et al., 2015; Rodell et al., 2009), drought monitoring (e.g., Houborg et al., 2012; Li et al., 2019) and forecasting (Getirana, Rodell, et al., 2020), agricultural planning and yield forecasting (Bernknopf et al., 2018), streamflow forecasting (Getirana, Jung, et al., 2020), flood vulnerability assessment and forecasting (Reager et al., 2014, 2015), and local sea level rise analysis (Caron et al., 2018; Han et al., 2019) (see Figure 1). The MCAT developed an online survey for applied science and non-science end users and potential future users, with questions that attempted to ascertain needs in terms of data type, continuity, spatial and temporal resolutions, accuracy, and timeliness. Based on the 87 survey responses and feedback from workshops, conference presentations, and interviews with stakeholders, high priority desires were determined to include improved timeliness (higher frequency, reduced latency) and increased spatial resolution relative to the standard GRACE and GRACE-FO products. In addition, potential new users, particularly in the government and industrial sectors, indicated that they would be unlikely to incorporate MC products into their operations if they lacked confidence that the products would continue to be available reliably and into the future. The information gathered by the MCAT had some bearing on the SATM, including the determination of the key variable for each of the 15 objectives, but it will likely have its biggest influence on decisions regarding a future MC science data system. For example, delivering a level-4 data assimilation product (e.g., Houborg et al., 2012; Li et al., 2019) could become a mission priority, as a majority of survey respondents preferred low latency (1 week or better) MC data with weekly or better temporal resolution and spatial resolution of $(25\ \mathrm{km})^2$ or better; data assimilation schemes are currently the only viable approach to achieve this desired spatial resolution. Further, the knowledge gained from MCAT engagement activities may be useful in targeting stakeholders for future MC data and information products across sectors that depend heavily on the availability of water, including irrigation, electricity generation, manufacturing, and the provision of municipal water.

Similar to the MCDO study, Pail et al. (2015) described consensus recommendations of an international panel of scientists for a next generation gravity mission. Another report, by a NASA-ESA interagency working group (NASA/ESA IGSWG, 2016), described observational targets for a future mass change observing system. The recommendations in the MC SATM differ from those in the two reports in some ways but are similar in others. Both reports began by defining the spatial and temporal scales of mass change signals from the various hydrological, cryospheric, oceanic, and solid earth sources of mass change. The DS, which was the basis for the MC SATM, largely agreed in its definition of those scales. Importantly, the MC SATM suggests a specific set of values for the three measurement variables (spatial and temporal resolution and accuracy) necessary for meaningful contribution of MC observations to each of the 15 MC-related objectives in the DS, considering also synergistic and complementary observations relevant for those objectives, while the recommendations of Pail et al. (2015) are meant to satisfy the needs of all user communities. NASA/ESA IGSWG (2016) is less prescriptive in its recommendations. The threshold requirements of Pail et al. (2015) are somewhat more aggressive than the MC baseline measurement parameters. For example, for a monthly terrestrial water storage (TWS) anomaly field, the former stipulates 5 mm accuracy at (400 km)^2 resolution, while 25 mm at (450 km)^2 is representative of the latter. The MC SATM goal parameters for the 15 objectives encompass a range of values that vary depending on the scientific objective. One specific example relates to objective S-3a focused on quantifying rates of sea level change and its driving processes, where the MC SATM goal parameter is explicitly stated in the DS as 10 mm accuracy at monthly timescales and (200 km)^2 resolution. This is in precise agreement with the target requirements described in Pail et al. (2015); however, we note there are also instances where the MC SATM goal parameters are more ambitious than the target requirements in Pail et al. (2015).

3. Architectures and Technology

3.1. Overview of Mass Change Architectures

The fundamental observable of any mass change architecture is the Earth's geopotential field, and how it changes over time. Inferring surface mass change from geopotential change is necessary to address most of the science questions in the MC SATM, and requires the application of well-understood loading theory (Wahr et al., 1998). Several of the Solid Earth science questions in the MC SATM (e.g., S-1b, S-3a, S-5a) can be more suitably addressed through assessing geopotential change, rather than surface mass change; however, for the purposes of this study, we consider the two synonymous as the required measurements are the same.

Spaceborne techniques for measuring global time variable gravity (i.e., mass change) have been of great interest to the science community for many decades. The most basic and oldest method is precise orbit determination (POD), which uses the observed positions and/or velocities of low Earth orbit (LEO) satellites to infer changes in the global gravity field. The earliest time variable gravity estimates were derived from satellite laser ranging (SLR) tracking data and were only able to recover several of the lowest degree (i.e., largest spatial wavelength) spherical harmonic coefficients (Cheng et al., 1997; Tapley et al., 1993). More recent studies have significantly expanded upon the number of estimated coefficients using GNSS tracking data to multiple LEO satellites (Richter et al., 2021; Teixeira Encarnação et al., 2020; Zhong et al., 2021). Doppler Orbitography and Radiopositioning Integrated by Satellite (DORIS) is another widely used POD technique that has also been applied for mass change studies (Cerri et al., 2013; Talpe et al., 2017). As early as the work of Wolff (1969), it was understood that a satellite-to-satellite tracking (SST) architecture with precise measurements of inter-satellite range changes between a pair of co-orbiting satellites promised to enhance the spatial resolution beyond what is possible with the POD approach. This general concept has, of course, since been successfully implemented as the GRACE (2002–2017) and GRACE-FO (2018-Present) mass change missions. Mass change would also be observable with spaceborne gravity gradiometers (GG), which can be realized in any single axis (or multiple axes) on a single satellite platform with a pair of separated accelerometers, if a certain threshold of instrument accuracy can be achieved. The POD, SST, and GG architectures and their associated technology options were investigated by the MCDO study team and are discussed in more detail throughout the remainder of this section. Additional architecture and technology details are provided in Text S1, Tables S2, S3, and S4 in Supporting Information S1.

For the sake of completeness, we briefly note several methods for measuring or inferring global time variable gravity signals that were excluded from the study. Despite their important contributions to measuring the higher spatial resolution components of the static gravity field, airborne and shipborne gravity gradiometers were not considered, as they lack the precision needed to recover temporal variability (Forsberg & Olesen, 2010; Sampietro et al., 2018). We also excluded from the study the investigation of GNSS ground stations (Argus et al., 2017; Borsa et al., 2014) and ground-based gravimeters (Breili & Rolstad, 2009; Güntner et al., 2017) due to the impracticality of deploying and maintaining the expansive network of instruments that would be required to observe mass change globally (among other challenges). Lastly, we note that the global gravity field can be inferred according to general relativity from very precise spaceborne clock measurements in combination with knowledge of the clock position and velocity (Müller et al., 2018). However, despite the substantial improvements of clock accuracy and stability in recent years, this concept is not currently under consideration given the stringent requirements on velocity accuracy and clock stability that is needed over short integration times.

3.2. Precise Orbit Determination

As summarized above, the POD approach derives mass change measurements from LEO satellite positions determined with GNSS, SLR, or DORIS tracking data. As the number of LEO satellites equipped with these precise tracking systems has increased, the ability to observe mass change signals from the POD method has improved, leading the study team to investigate the potential performance of a dedicated mass change POD constellation. The architecture trade space is defined by the number and arrangement of satellites, while the technology trade space consists of the tracking system, attitude determination system, and the possible inclusion of an accelerometer for measuring the non-gravitational forces. Our simulation study began with an overly optimistic implementation in order to assess the "ceiling" of POD performance. We assumed that all satellites are flown at low altitudes, where each is equipped with a geodetic-quality GNSS receiver, and an accelerometer and attitude determination system with performance equivalent to that flown on GRACE-FO. We considered orbit configurations with both single and multi-plane arrangements that optimize the spatiotemporal sampling, and simulated constellation sizes of 24, 48, and 96 satellites. We also simulated a scenario with a constellation of co-orbiting satellite pairs (similar to GRACE), where kinematic baseline ranges were computed and incorporated as observation data. This POD-based approach is motivated by the superior accuracy of the baseline ranges (millimeter level) relative to the absolute positions (centimeter level) due to the cancellation of common errors (Guo & Zhao, 2019; Teixeira Encarnação et al., 2020). Even for the largest and most overly optimistic POD constellation scenario, the computed science value was far below an acceptable level for the team to consider further study of the POD option (Figure 2). This result is effectively confirmed by the analysis of Zhong et al. (2021), which also concludes that a sizable constellation of GNSS-equipped LEO satellites does not approach the spatial resolution of a dedicated SST mission, such as GRACE.

3.3. Satellite-to-Satellite Tracking

Given the long program of record of the GRACE missions, extensive work by the science and engineering communities to study and advance SST architectures and technologies pre-dates our study. Our team's thorough review of the scientific literature, mission proposals, and technology development efforts from the past few decades was essential to identify the SST trade space we investigated. The SST architecture trade space included: (a) single in-line pair (two satellites in the same orbital plane separated in the along-track direction, similar to the GRACE and GRACE-FO architectures); (b) single pendulum pair (two satellites with small differences in the right ascension of the ascending node and mean anomaly, where an opening angle specifices the angle at the equatorial crossing between the equator and the line of sight between the two satellites; this formation results in a combination of north–south measurements (maximum at poles) and east–west measurements (maximum at the equator), as the satellites progress in their relative orbits (Sharifi et al., 2007); (c) in-line pair plus a third trailing satellite that forms a pendulum; (d) two in-line pairs, commonly referred to as a Bender formation (one pair in a polar orbit, one pair in a lower inclined orbit, typically between 65° and 75°) (Bender et al., 2008); (e) LEO-MEO (low Earth orbit satellite(s) ranging between medium Earth orbit satellite(s) (Hauk & Pail, 2019)); and (f) SmallSat/CubeSat constellation of satellite pairs performing SST. Cartwheel and helix configurations, which have been previously studied (Elsaka, 2014; Wiese et al., 2009), were omitted given their substantial complexity and limited performance benefit relative to the SST configurations we considered.

SST-relevant technology development efforts can be grouped into two categories: (a) advancements to existing technologies that would benefit a single or dual in-line pair architecture like GRACE/GRACE-FO; and (b) new technologies that enable new architecture configurations. The first category is primarily focused on improving the performance or redundancy of the inter-satellite ranging and accelerometer instruments, and we note that the attitude determination system is an important supporting technology as well. The second category of development efforts includes technologies required to: fly at a lower altitude and/or perform regular orbit maintenance (e.g., electric propulsion for a drag compensation system); implement a pendulum architecture (e.g., frequency comb and laser chronometer); implement a LEO-MEO architecture (e.g., laser chronometer); reduce the size, weight, and power (SWaP) for all relevant technologies for a cost-effective multi-platform SmallSat/CubeSat SST constellation (e.g., inter-satellite ranging system and accelerometers). Given its importance to both technology development categories, and the extensive development work that is underway, much of our study focused on the science value impacts of the inter-satellite ranging and accelerometer options.

The simulated recovery of mass change signals was performed for the large suite of SST architecture and technology options briefly summarized above and captured in Tables S2, S3, and S4 in Supporting Information S1. We worked closely with the instrument developers to incorporate proper error budgets into the simulations and to capture the expected SWaP, technology readiness level (TRL), and planned development schedules for each technology. For the inter-satellite ranging technologies, performance simulations were executed for the GRACE-FO microwave interferometer (MWI) and laser ranging interferometer (LRI), the reduced-SWaP K-/V-band ranging system in development at GeoOptics, Inc., the reduced-SWaP μNPRO in development at NASA GSFC, and the laser chronometer in development at CNES. The optical frequency comb in development at Ball Aerospace and laser metrology interferometer in development at ESA, are expected to have performance similar to the LRI, so separate error budgets were not needed for those technologies. We note here that the LRI was a successful technology demonstration instrument on GRACE-FO (TRL 9), and has provided measurement performance significantly exceeding the MWI (Abich et al., 2019), while not compromising the quality of the mass change estimates (Peidou et al., 2022; Pie et al., 2021). Our study team worked with the JPL engineers to outline the set of well-defined standard engineering steps, and a development schedule required for the LRI to be flown as a primary instrument.

For accelerometers we considered a range of current and developing technologies at ONERA (vendor for GRACE/GRACE-FO), as well as the Simplified LISA Pathfinder Gravitational Reference Sensor (S-GRS) (Davila Alvarez et al., 2021) and compact optomechanical accelerometers (Hines et al., 2020). Given the current TRL of the various development efforts, the team recommended use of an ONERA electrostatic accelerometer for the next mass change mission. We note that different design specifications can be levied depending on the selected architecture, altitude, and inclusion of a drag compensation system. The study team identified the value of considering both the S-GRS and optomechanical technologies as potential technology demonstrators for the next mission. The S-GRS promises several orders of magnitude improvement in performance relative to the GRACE-FO accelerometers, which are presently the largest source of measurement system error. The significant reduction in SWaP of the optomechanical device could facilitate redundancy with minimal impact on the spacecraft design, while also advancing efforts to miniaturize all SST-relevant technologies.

3.4. Gravity Gradiometers

The use of spaceborne gravity gradiometers for measuring the static gravity field was successfully implemented by the GOCE mission (2009–2013) for which six electrostatic accelerometers were arranged to form gravity gradiometers along each of the three orthogonal axes (Bouman & Fuchs, 2012). These accelerometers lacked the required precision for estimating temporal gravity changes at monthly time scales; however the information provided on the static gravity field is relevant for mass change observations as it can be used to define a reference gravity field used in the data processing and numerical simulations (Section 4). The ongoing development of atomic interferometer gravity gradiometer (AIGG) technology promises to address the performance limitation of the GOCE gravity gradiometers and enable accurate mass change measurements from a single satellite. This emerging technology captures the influence of the gravity field on a cloud of atoms (Carraz et al., 2014), and our team's simulations demonstrated high science value for a single AIGG instrument oriented in the radial direction. Multiple GG instruments oriented in the other orthogonal directions would add information to the solution as well, and a hybrid single pair SST architecture equipped with a precise GG would improve performance relative to the SST-only configuration. The mass change study advanced this technology through instrument and mission design lab studies conducted at NASA GSFC in collaboration with engineers at AOSense, Inc. Despite its promise of high science value, the GG option was not recommended for the next mass change mission due to the uncertain AIGG development schedule.

4. Architecture Assessment Process

Numerical simulations are used to assess the performance of the architectures and technologies (Section 3) relative to the measurement parameters identified in the SATM (Section 2). Such simulations have been widely used in the literature to perform similar assessments (Elsaka, 2014; Flechtner et al., 2016; Hauk & Wiese, 2020; Loomis et al., 2012; Wiese et al., 2012), with software and processes that rely on and mimic the processing of GRACE and GRACE-FO data. Two types of simulations are performed: (a) those that include both measurement system error and temporal aliasing error, the latter of which is well understood to be a limiting source of error

10.1029/2022EA002311

for SST satellite gravimetry missions (Flechtner et al., 2016; Han et al., 2004; Thompson et al., 2004; Wiese et al., 2012), and (b) those that include only measurement system error. Simulations of type 1 are used to derive a science value score for each architecture, which provides a best estimate of the expected quality of the mass change data products. Simulations of type 2 are used to derive a measurement system value, which represents the best performance that could be achieved if temporal aliasing error is mitigated in the future via either improved models of high frequency mass variations or improved data processing strategies.

Details on the numerical simulation process are provided in Text S2 and Table S5 in Supporting Information S1. The simulations rely on the creation of a truth run where simulated measurements are created using realistic force models to define the flight environment; these models include the mass change signals of interest. A nominal run is then performed where perturbations are introduced relative to the truth run; these perturbations include errors in background force models (i.e., temporal aliasing error) and realistic errors on the measurement system. Measurement system errors consist of inter-satellite ranging, accelerometer, attitude, and absolute position measurement errors, and are derived from multiple sources (Tables S3 and S4 in Supporting Information S1). For instruments with heritage from GRACE-FO (LRI, MWI, GRACE-FO accelerometer, GNSS, attitude knowledge), error spectra are either derived from GRACE-FO flight data where available, or best estimates of instrument performance prior to the launch of GRACE-FO. For instruments in development with little to no flight heritage, a characterization of the errors across the relevant frequency spectra has been provided by the developer, and those spectra are used to derive the instrument errors introduced in the numerical simulations. Residuals are created by differencing simulated measurements from the truth and nominal runs and these residuals are used to estimate the truth environment in the presence of the errors in a large linear least squares inversion process. Errors are quantified by differencing the estimated gravity field from the truth gravity field using 1 month (i.e., the targeted temporal resolution of each objective in the SATM) of simulated data. These errors are then mapped to a range of spatial scales (110–1,000 km) by smoothing the signals of interest using a Gaussian filter at the relevant spatial scales, similar to how errors in GRACE-FO have been quantified (Landerer et al., 2020).

Equation 1 is used to derive a science value (SV) for each architecture (α), which scores its ability to achieve the baseline measurement parameters in the SATM (Table S1), and thus, be responsive to the DS science objectives.

$$\mathrm{SV}(\alpha) = \frac{\sum_{n=1}^{15} W_n P_n(\alpha)}{\sum_{n=1}^{15} W_n};\ \text{if} \begin{cases} \text{Key Variable}_n = \text{AC};\ P_n(\alpha) = \dfrac{\mathrm{AC}_n}{\mathrm{AC}(\alpha)|_{\mathrm{SR}_n \mathrm{TR}_n}} \\ \text{Key Variable}_n = \text{SR};\ P_n(\alpha) = \dfrac{\mathrm{SR}_n}{\mathrm{SR}(\alpha)|_{\mathrm{AC}_n \mathrm{TR}_n}} \end{cases} \tag{1}$$

Here, n represents a science objective in the SATM, and W_n is the weight of that objective, defined as the Importance multiplied by the Utility. Numerical values of (*0.33; 0.67; 1*) are prescribed for Importance Scores of (*Important; Very Important; Most Important*) and Utility scores of (*Low, Medium, High*), respectively. A Very Low Utility score is prescribed to be 0.1. Study results were found to be independent of the choice of numerical value for the weight. P_n represents the performance of the architecture, which is dependent upon the key variable, as defined in the SATM, of either accuracy (AC), spatial resolution (SR), or temporal resolution (TR). In essence, the performance of an architecture is assessed by quantifying error across space and time (similar to Hauk & Wiese, 2020), and then scored dependent upon how well the key variable can be estimated in that domain. For example, science objective H-1a lists accuracy as the key variable (Table S1) with a target of 10 mm; therefore to evaluable $P_{H\text{-}1a}$ for a given architecture, one would divide the targeted accuracy (10 mm) by the accuracy the particular architecture can achieve (derived from numerical simulation results) at the targeted spatial (1,000 km)^2 and temporal (monthly) resolutions for H-1a. We note that since only one science objective (H-4c; Important) had key variable equal to temporal resolution, spatial resolution was assigned as a secondary key variable to this objective to save on computing resources. The denominator of Equation 1 normalizes the science value against the sum of the weights. Since the SATM baseline measurement parameters were constructed to represent performance of the program of record, this in essence means that science value = 1 represents architecture performance that is equivalent to the program of record; science value <1 represents degradation relative to the program of record; and science value >1 represents improvements relative to the program of record. Science value = 3, for example, can be interpreted as improvements in some combination of resolution/accuracy by a factor of 3 relative to the program of record.

10.1029/2022EA002311

5. Value Framework

The value framework provides a mechanism for objectively comparing and discriminating between the candidate observing systems identified by the MCDO study team. It is the basis for the assessment and evaluation processes applied by the study team to identify and to make a recommendation to NASA's Earth Science Division on which candidate observing system architectures should be further studied for pre-formulation activities. The value framework must allow multiple candidate observing systems to be compared, including aspects of effectiveness and affordability, as well as other factors such as compatibility with potential international partnerships and existing NASA policies. The effectiveness of candidate observing systems was primarily measured by science value and risk, whereas the affordability was measured by estimates of cost, schedule, and budget availability.

In addition to these traditional areas of assessment, the value framework also considered the probability of providing continuity between the next MC observing system and the program of record, since this was a key goal for MC as expressed in the DS. The likelihood of maintaining continuity is driven by both the expected development cycle for the next MC observing system, and the expected end of life of GRACE-FO. The value framework considers both factors by examining stochastic estimates for the development schedules for each candidate observing system and comparing against the expected range of end of life dates for the GRACE-FO mission. To estimate the end of life date for GRACE-FO, we considered two triggering mechanisms: an on-orbit failure of the spacecraft leading to loss of science, and the gradual degradation of the GRACE-FO orbital altitude due to atmospheric drag. To understand the likelihood of a failure triggering end of life, the team leveraged historical spacecraft reliability data for similar class missions (Ferrone et al., 2019) and derived a Weibull distribution to represent the probability of a failure as a function of mission duration. To understand the likelihood of orbital altitude degradation triggering end of life, the team leveraged predictions for solar cycle 25 and 26 (Pesnell & Schatten, 2018) to perform stochastic orbit lifetime analysis using initial spacecraft conditions based on the GRACE-FO mission parameters. Combining the historical spacecraft reliability and orbit lifetime estimates allowed the estimation of a range of dates for the GRACE-FO end of life (Figure S1 in Supporting Information S1) which could then be compared to the MC candidate observing system development schedules and launch readiness estimates to understand the likelihood of maintaining continuity between GRACE-FO and each of the MC candidate observing systems. Figure S1 in Supporting Information S1 shows that the GRACE-FO end of life is more likely to be triggered by on-orbit failure than degradation of the orbital altitude, with estimates of spacecraft reliability of 70% and 50% occurring in 2025 and 2028, respectively.

6. Results

One useful evaluation metric in the value framework process is an assessment of science value versus implementation cost (Figure 2) for all architecture classes. It was found that POD architectures are not capable of meeting the baseline measurement parameters (science value = 0.1), and do not scale well with increasing numbers of elements (24 elements increasing to 96 elements increases science value from 0.08 to 0.12) even when the most optimistic assumptions on instrument performance and orbit geometry are used (see Section 3.2); thus, POD was eliminated from further consideration. It also became apparent that while GG architectures provide the potential for high science return (science value up to 3.5), the relatively low technical maturity and unclear plans for further maturation of GG technologies made this an unfavorable candidate for further study as an observing system that could be implemented this decade. A GG architecture for MC would significantly jeopardize the ability to provide continuity with GRACE-FO relative to SST architectures that were studied.

Significant challenges were also identified for two subsets of SST architectures: the LEO-MEO architecture and a constellation of SmallSats/CubSats. Challenges for the LEO-MEO architecture included operational constraints on the inter-satellite ranging systems and restrictions on allowable laser power due to concerns of potentially lasing other space assets. Since the LEO-MEO SST architectures did not provide science performance increases above heritage single in-line pair architectures (science value = 1.12 for a 4-satellite LEO-MEO 1 architecture (Table S2 in Supporting Information S1) versus science value = 1.14 for a single in-line pair architecture), while facing significant challenges, they were also eliminated from further consideration. The viability of a constellation of SmallSats/CubeSats was studied through a dedicated Team X exercise conducted at JPL. Team X consists of a multi-disciplinary team of engineers that utilizes concurrent engineering methodologies to rapidly design, analyze, and evaluate mission concept designs. The Team X study goal was to determine whether an SST architecture exists within the same mission risk classification as GRACE-FO (i.e., Class C), that meets the baseline measurement parameters of the SATM while satisfying the MC cost target documented in the DS, by leveraging smaller, less mature technologies and components. The findings of the Team X study were that the form factor of the spacecraft bus could be reduced relative to what has been flown on the program of record; however, the cost target in the DS was still exceeded in addition to significantly increasing mission risk; hence, the SmallSat architecture concept was eliminated from further consideration as the next MC observing system.

The remaining observing systems in the tradespace were all SST architectures in different configurations, including single in-line pairs, pendulum pairs, a 3-satellite architecture combining an in-line pair with a pendulum satellite, and two pair Bender configurations (one polar pair coupled with a pair at a lower inclination). Each of those configurations included variations in the orbit altitude and instrumentation, including different ranging system and accelerometer options (Tables S3 and S4 in Supporting Information S1). Based on the readiness of technologies associated with each configuration, further reductions in the tradespace were made, removing architectures utilizing the S-GRS, HybridSTAR, and optomechanical inertial sensor, as those accelerometer technologies were unlikely to be ready for flight mission implementation in time for the next MC observing system, and were better suited as potential technology demonstrator candidates. The LRI was selected as the best option for the inter-satellite ranging instrument for in-line pair observing system components due to its successful demonstration on GRACE-FO and superior measurement system value relative to the MWI. The addition of an optical frequency comb to the LRI, along with the laser chronometer were retained as inter-satellite ranging technologies for observing system components that require a pendulum formation. After pruning based on TRL and measurement system value, ten distinct architectures remained (Figure 2), which are described in more detail in Table 1.

Examination of the remaining tradespace (Figure 2, right) provides some initial observations: (a) all of the remaining architectures are capable of meeting the baseline science objectives, while none meet the DS cost target, and (b) within an architecture type, variation in cost is primarily driven by technologies and payloads while variation in science value is primarily driven by orbital characteristics. The subset of Bender architectures has the largest variations in science value ranging from 2.75 (D1; both pairs at 500 km) to 4.15 (D4; both pairs at 350 km). The other two Bender configurations (D2, D3) mix high and low altitude pairs, and show a significant difference in science value, with the higher performing option placing the lower altitude pair in the inclined orbit (D3). An important result of the study is that the most significant contribution to science value from the Bender architecture comes from placing the inclined pair in a lower altitude. The altitude of the polar pair can be regarded as a secondary design variable, where science value shows only a modest decrease from 4.15 to 3.85 due to raising the altitude of the polar pair from 350 km (D4) to 500 km (D3), and also dropping the need for a drag compensation system that is required when flying at lower altitudes. It is additionally worth noting that, assuming no flight system failures, a satellite at 350 km with a drag compensation system is limited in lifetime by the amount of consumables onboard; after the consumables are exhausted, the satellite will re-enter Earth's atmosphere within the timeframe of several months. Conversely, a satellite at 500 km without a drag compensation system is limited

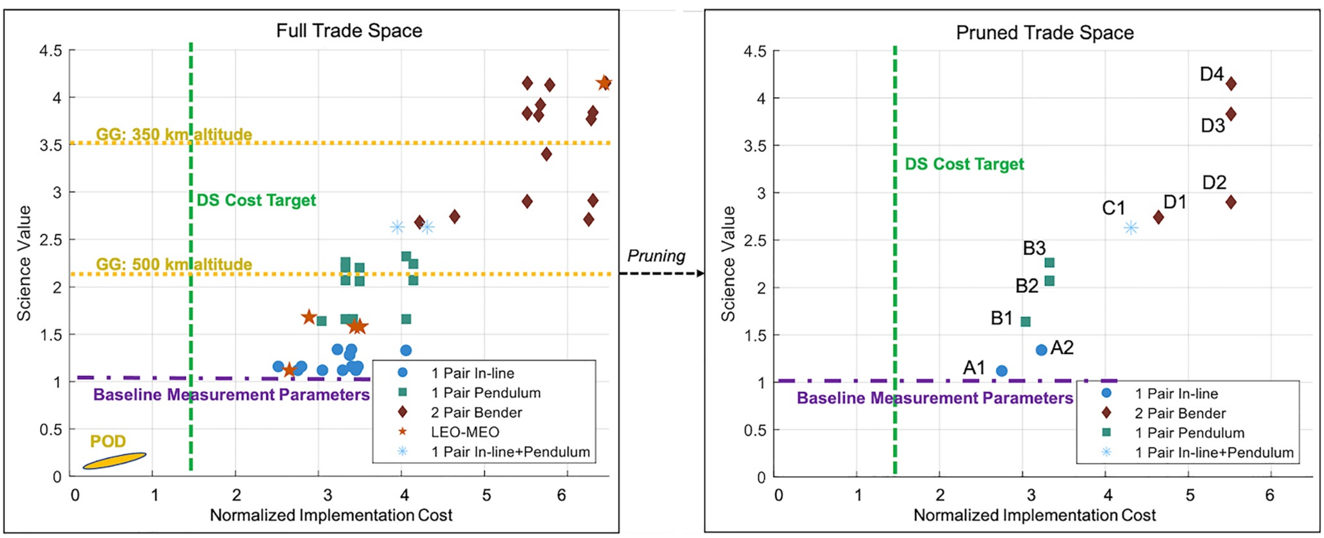

Figure 2. Science value versus normalized implementation cost of the full trade space of architectures (left), and the pruned trade space (right).

Table 1
Observing System Characteristics of Remaining Architectures After an Initial Pruning Stage in the Value Framework

Observing system					
Plain language description	ID	Number of platforms	Altitude (km) (Opening angle)	Payload	Probability of overlap
Single In-line Pair: High Altitude ***[Baseline]***	A1	2	500	1	50%
Single In-line Pair; Low Altitude	A2	2	350	2	35%
Single Pendulum Pair; High Altitude, Large Opening Angle	B1	2	500 (OA = 45°)	3	40%
Single Pendulum Pair; Low Altitude, Small Opening Angle	B2	2	350 (OA = 15°)	4	35%
Single Pendulum Pair; Low Altitude, Large Opening Angle	B3	2	350 (OA = 45°)	4	35%
In-line Pair + Pendulum; High Altitude, Large Opening Angle ***[Enhancing]***	C1	3	500 (OA = 45°)	In-line:1; Pend: 3	In-line: 50% Pend: 40%
Bender: High-High ***[Enhancing]***	D1	4	PP: 500; IP: 500	PP: 1; IP: 1	50%
Bender: Low-High	D2	4	PP: 350; IP: 500	PP: 2; IP: 1	PP: 35% IP: 50%
Bender: High-Low ***[Enhancing]***	D3	4	PP: 500; IP: 350	PP: 1; IP: 2	PP: 50%; IP: 35%
Bender: Low-Low	D4	4	PP: 350; IP: 350	PP: 2; IP: 2	35%

Note. Multi-element observing systems include characteristics for the polar pair (PP), inclined pair (IP), in-line pair, and pendulum (Pend) satellite separately, and opening angles (OA) for the pendulum formation are specified. Payload identifications are as follows: 1 = LRI + ONERA GRACE-FO accelerometer; 2 = LRI + ONERA MicroSTAR + Drag Compensation (DC); 3 = Laser Chronometer (LC) + ONERA MicroSTAR-Prime; 4 = LC + ONERA MicroSTAR + DC. The probability of overlap represents the expected reliability of GRACE-FO at the 50th percentile launch readiness date for each respective architecture.

in lifetime by natural altitude degradation due to atmospheric drag (lifetime has potential to be >15 years depending on solar activity), leading to a potentially longer lifetime than a satellite at lower altitude.

A significant discriminator among architectures in Table 1 is their probability to provide overlap with GRACE-FO (see Figure S1 in Supporting Information S1 for estimates of GRACE-FO end of life). Architectures A1, C1, D1, and D3 have the highest likelihood of providing continuity with the program of record (50% probability at the 50th percentile launch readiness date) by launching a new pair of satellites in a polar orbit. Each of these architectures include one component that is a single in-line pair flown in a polar orbit at 500 km altitude (Architecture A1), leveraging heritage technology from GRACE and GRACE-FO, leading to the shortest expected development schedule and lowest cost; because of this, Architecture A1 (Table 1) is labeled as the *Baseline* observing system. Architectures C1, D1, and D3 are multi-element observing systems that include Architecture A1 as one of the elements. Each of these are similar to architectures under consideration by international space agencies, and provide potential international partnership opportunities for consideration by NASA. These architectures (labeled as *Enhancing* in Table 1), could be implemented in a phased approach with the polar pair being developed and launched first to minimize the likelihood of a data gap with the program of record, and the remaining elements of the observing system launched 1–2 years later for full system completion. This evolvable implementation approach offers programmatic flexibility towards satisfying the MC baseline measurement parameters with a low risk posture using Architecture A1, while simultaneously investing in technological advancements necessary for implementing Architectures C1 and D3 (possibly in collaboration with other international space agencies) to gain significant increases in science value relative to the program of record. It is worth noting that the highest performing architecture identified (D3) satisfies one of the goal measurement parameters in the MC SATM, while nearly satisfying several more.

7. Conclusions

In this manuscript, we provide a high level overview and main results of the MCDO study. The objective of the study was to identify a small subset of high value observing systems for further study that could be implemented within the current decade that are responsive to the scientific objectives of the DS. The study framework included generation of a MC SATM, the identification of three architecture classes for measuring mass change, the use of a numerical simulation framework to quantify architecture performance relative to the SATM and derive science value scores, and a larger value framework process to provide a recommendation to NASA. The value framework process considered multiple aspects of each potential architecture to understand and quantify value.

These attributes included science value, cost, technical risk, international partnership opportunities, and schedule, including the likelihood of overlap with GRACE-FO.

The primary outcome of the study is the recommendation that an SST architecture be implemented for the MC observing system. A single in-line pair architecture similar in nature to both GRACE and GRACE-FO was identified as the lowest cost architecture capable of meeting the baseline measurement parameters, while also having the highest probability of providing continuity with GRACE-FO (50% probability of providing overlap); as such, this architecture is identified as the Baseline observing system. Two enhancing elements that can potentially leverage international partnership opportunities were identified to improve the science value relative to the Baseline observing system. The first is the addition of a second pair of satellites inserted into a complementary inclined orbital plane, and the second is the addition of a third satellite to the Baseline observing system that performs a relative pendulum motion. Both enhancing elements have potential to be added modularly as soon as 1–2 years after launch of the Baseline observing system to complete the final observing system. The high value observing systems (Baseline + Enhancing) recommended in this manuscript are now under study in more depth by NASA and potential international partners to arrive at a final mass change observing system for implementation as a core component of the Earth System Observatory.

Data Availability Statement

Data generated in the manuscript are based on output from model simulations. Appropriate model references are provided throughout the manuscript, along with configuration information for the model runs. Numerical simulations to recover the gravity field were run at both JPL (using the MIRAGE software suite) and at GSFC (using the GEODYN software suite); both software packages are restricted from being shared due to proprietary concerns and Export Administration Regulations from the U.S. government. However, appropriate descriptions of the numerical simulation process and setup are provided in the Supplement, so users may replicate the process with their own software tools. Spacecraft reliability data used are found in Ferrone et al. (2019), while orbit lifetime analysis is driven by input from Pesnell and Schatten, (2018). Costing and schedule development models used are proprietary.

Acknowledgments
A portion of this work was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration (80NM0018D0004).

References

Abich, K., Abramovici, A., Amparan, B., Baatzsch, A., Okihiro, B. B., Barr, D. C., et al. (2019). In-orbit performance of the GRACE follow-on laser ranging interferometer. *Physical Review Letters*, *123*(3), 031101. https://doi.org/10.1103/PhysRevLett.123.031101

Argus, D. F., Landerer, F. W., Wiese, D. N., Martens, H. R., Fu, Y., Famiglietti, J. S., et al. (2017). Sustained water loss in California's mountain ranges during severe drought from 2012 to 2015 inferred from GPS. *Journal of Geophysical Research: Solid Earth*, *122*(12), 10559–10585. https://doi.org/10.1002/2017JB014424

Bender, P. L., Wiese, D. N., & Nerem, R. S. (2008). A possible dual-GRACE mission with 90 degree and 63 degree inclination orbits. In *Proceedings of the third international symposium on formation flying, missions and technologies. 23-25 April 2008* (pp. 1–6).

Bernknopf, R., Brookshire, D., Kuwayama, Y., Macauley, M., Rodell, M., Thompson, A., et al. (2018). The value of remotely sensed information: The case of a GRACE-enhanced drought severity index. *Weather, Climate, and Society*, *10*(1), 187–203. https://doi.org/10.1175/WCAS-D-16-0044.1

Borsa, A. A., Agnes, D. C., & Cayan, D. R. (2014). Ongoing drought-induced uplift in the western United States. *Science*, *345*(6204), 1587–1590. https://doi.org/10.1126/science.1260279

Bouman, J., & Fuchs, M. J. (2012). GOCE gravity gradients versus global gravity field models. *Geophysical Journal International*, *189*(2), 846–850. https://doi.org/10.1111/j.1365-246X.2012.05428.x

Breili, K., & Rolstad, C. (2009). Ground-based gravimetry for measuring small spatial-scale mass changes on glaciers. *Annals of Glaciology*, *50*(50), 141–147. https://doi.org/10.3189/172756409787769717

Caron, L., Ivins, E. R., Larour, E., Adhikari, S., Nilsson, J., & Blewitt, G. (2018). GIA model statistics for GRACE hydrology, cryosphere, and ocean science. *Geophysical Research Letters*, *45*(5), 2203–2212. https://doi.org/10.1002/2017GL076644

Carraz, O., Siemes, C., Massotti, L., Haagmans, R., & Silvestrin, P. (2014). A spaceborne gravity gradiometer concept based on cold atom interferometers for measuring Earth's gravity field. *Microgravity Science and Technology*, *26*, 139–145. https://doi.org/10.1007/s12217-014-9385-xsatm

Cerri, L., Lemoine, J. M., Mercier, F., Zelensky, N. P., & Lemoine, F. G. R. (2013). DORIS-based point mascons for the long term stability of precise orbit solutions. *Advances in Space Research*, *52*(3), 466–476. https://doi.org/10.1016/j.asr.2013.03.023

Cheng, M. K., Shum, C. K., & Tapley, B. D. (1997). Determination of long-term changes in the Earth's gravity field from satellite laser ranging observations. *Journal of Geophysical Research*, *102*(B10), 22377–22390. https://doi.org/10.1029/97JB01740

Davila Alvarez, A., Knudtson, A., Patel, U., Gleason, J., Hollis, H., Sanjuan, J., et al. (2021). A simplified gravitational reference sensor for satellite geodesy. *Journal of Geodesy*. Preprint available: arXiv:2107.08545 [physics.ins-det].

Elsaka, B. (2014). Feasible multiple satellite mission scenarios flying in a constellation for refinement of the gravity field recovery. *International Journal of Geosciences*, *5*(03), 267–273. https://doi.org/10.4236/ijg.2014.53027

Famiglietti, J. S., Lo, M., Ho, S. L., Bethune, J., Anderson, K. J., Syed, T. H., et al. (2011). Satellites measure recent rates of groundwater depletion in California's Central Valley. *Geophysical Research Letters*, *38*(3), 2010GL046442. https://doi.org/10.1029/2010GL046442

Ferrone, K., Tran, T., Cavanaugh, M., & Brown, N. (2019). 2019 Satellite lifetime study, Aerospace report No. TOR-2019-02620.

Flechtner, F., Neumayer, K.-H., Dahle, C., Dobslaw, H., Fagiolini, E., Raimondo, J.-C., & Guntner, A. (2016). What can be expected from the GRACE-FO laser ranging interferometer for Earth science applications? *Surveys in Geophysics*, *37*(2), 453–470. https://doi.org/10.1007/s10712-015-9338-y

Forsberg, R., & Olesen, A. V. (2010). Airborne gravity field determination. In G.Xu (Ed.) *Sciences of Geodesy, I*. Springer. https://doi.org/10.1007/978-3-642-11741-1_3

Getirana, A., Jung, H. C., Arsenault, K., Shukla, S., Kumar, S., Peters-Lidard, C., et al. (2020). Satellite gravimetry improves seasonal streamflow forecast initialization in Africa. *Water Resources Research*, *56*(2), e2019WR026259. https://doi.org/10.1029/2019WR026259

Getirana, A., Rodell, M., Kumar, S., Beaudoing, H. K., Arsenault, K., Zaitchik, B., et al. (2020). GRACE improves seasonal groundwater forecast initialization over the United States. *Journal of Hydrometeorology*, *21*(1), 59–71. https://doi.org/10.1175/JHM-D-19-0096.1

Güntner, A., Reich, M., Mikolaj, M., Creutzfeldt, B., Schroeder, S., & Wziontek, H. (2017). Landscape-scale water balance monitoring with an iGrav superconducting gravimeter in a field enclosure. *Hydrology and Earth System Sciences*, *21*(6), 3167–3182. https://doi.org/10.5194/hess-21-3167-2017

Guo, X., & Zhao, Q. (2019). A new approach to Earth's gravity field modeling using GPS-derived kinematic orbits and baselines. *Remote Sensing*, *11*(14), 1728. https://doi.org/10.3390/rs11141728

Han, S.-C., Jekeli, C., & Shum, C. K. (2004). Time-variable aliasing effects of ocean tides, atmosphere, and continental water mass on monthly mean GRACE gravity field. *Journal of Geophysical Research*, *109*(B4). https://doi.org/10.1029/2003JB002501

Han, S. C., Sauber, J., Pollitz, F., & Ray, R. (2019). Sea level rise in the Samoan Islands escalated by viscoelastic relaxation after the 2009 Samoa-Tonga earthquake. *Journal of Geophysical Research: Solid Earth*, *124*(4), 4142–4156. https://doi.org/10.1029/2018JB017110

Hauk, M., & Pail, R. (2019). Gravity field recovery using high-precision, high-low inter-satellite links. *Remote Sensing*, *11*(5), 537. https://doi.org/10.3390/rs11050537

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

Hines, A., Richardson, L., Wisniewski, H., & Guzman, F. (2020). Optomechanical inertial sensors. *Applied Optics*, *59*(22), G167–G174. https://doi.org/10.1364/AO.393061

Houborg, R., Rodell, M., Li, B., Reichle, R., & Zaitchik, B. F. (2012). Drought indicators based on model-assimilated Gravity Recovery and Climate Experiment (GRACE) terrestrial water storage observations. *Water Resources Research*, *48*(7), W07525. https://doi.org/10.1029/2011WR011291

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

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

Loomis, B. D., Nerem, R. S., & Luthcke, S. B. (2012). Simulation study of a follow-on gravity mission to GRACE. *Journal of Geodesy*, *86*(5), 319–335. https://doi.org/10.1007/s00190-011-0521-8

Müller, J., Dirkx, D., Kopeikin, S. M., Lion, G., Panet, I., Petit, G., & Visser, P. N. A. M. (2018). High performance clocks and gravity field determination. *Space Science Reviews*, *214*(1), 5. https://doi.org/10.1007/s11214-017-0431-z

NASA/ESA Interagency Gravity Science Working Group (IGSWG). (2016). Towards a sustained observing system for mass transport to understand global change and to benefit society. Doc. Nr: TUD-IGSWG-2016-01.

National Academies of Sciences, Engineering, and Medicine. (2018). *Thriving on our changing planet: A decadal strategy for earth observation from space*. The National Academies Press. https://doi.org/10.17226/24938

Pail, R., Bingham, R., Braitenberg, C., Dobslaw, H., Eicker, A., Guntner, A., et al. (2015). Science and user needs for observing global mass transport to understand global change and to benefit society. *Surveys in Geophysics*, *36*, 743–772. https://doi.org/10.1007/s10712-015-9348-9

Peidou, A., Landerer, F. W., Wiese, D. N., Ellmer, M., Fahnestock, E., McCullough, C., et al. (2022). Spatiotemporal characterization of geophysical signal detection capabilities of GRACE-FO. *Geophysical Research Letters*, *49*(1), e2021GL095157. https://doi.org/10.1029/2021GL095157

Pesnell, W. D., & Schatten, K. H. (2018). An early prediction of the amplitude of solar cycle 25. *Solar Physics*, *293*(7), 112. https://doi.org/10.1007/s11207-018-1330-5

Pie, N., Bettadpur, S. V., Tamisiea, M., Krichman, B., Save, H., Poole, S., et al. (2021). Time variable Earth gravity field models from the first spaceborne laser ranging interferometer. *Journal of Geophysical Research: Solid Earth*, *126*(12), e2021JB022392. https://doi.org/10.1029/2021JB022392

Reager, J. T., Thomas, A. C., Sproles, E. A., Rodell, M., Beaudoing, H. K., Li, B., & Famiglietti, J. S. (2015). Assimilation of GRACE terrestrial water storage observations into a land surface model for the assessment of regional flood potential. *Remote Sensing*, *7*(11), 14663–14679. https://doi.org/10.3390/rs71114663

Reager, J. T., Thomas, B. F., & Famiglietti, J. S. (2014). River basin flood potential inferred using GRACE gravity observations at several months lead time. *Nature Geoscience*, *7*(8), 588–592. https://doi.org/10.1038/NGEO2203

Richey, A. S., Thomas, B. F., Lo, M.-H., Reager, J. T., Famiglietti, J. S., Voss, K., et al. (2015). Quantifying renewable groundwater stress with GRACE. *Water Resources Research*, *51*(7), 5238. https://doi.org/10.1002/2015WR017349

Richter, H. M. P., Lück, C., Klos, A., Sideris, M. G., Rangelova, E., & Kusche, J. (2021). Reconstructing GRACE-type time-variable gravity from the Swarm satellites. *Scientific Reports*, *11*(1), 1117. https://doi.org/10.1038/s41598-020-80752-w

Rodell, M., Velicogna, I., & Famiglietti, J. S. (2009). Satellite-based estimates of groundwater depletion in India. *Nature*, *460*(7258), 999–1002. https://doi.org/10.1038/nature08238

Sampietro, D., Mansi, A. H., & Capponi, M. (2018). A new tool for airborne gravimetry survey simulation. *Geosciences*, *8*(8), 292. https://doi.org/10.3390/geosciences8080292

Sharifi, M., Sneeuw, N., & Keller, W. (2007). Gravity recovery capability of four generic satellite formations. In *1st international symposium of the international gravity field service* (Vol. *18*, pp. 211–216). ISSN 1300-5790.

Talpe, M. J., Nerem, R. S., Forootan, E., Schmidt, M., Lemoine, F. G., Enderlin, E. M., & Landerer, F. W. (2017). Ice mass change in Greenland and Antarctica between 1993 and 2013 from satellite gravity measurements. *Journal of Geodesy*, *91*(11), 1283–1298. https://doi.org/10.1007/s00190-017-1025-y

Tapley, B. D., Schutz, B. E., Eanes, R. J., Ries, J. C., & Watkins, M. M. (1993). Lageos laser ranging contributions to geodynamics, geodesy, and orbital dynamics. In D. E. Smith, & D. L. Turcotte (Eds.), *Contributions of space geodesy to geodynamics: Earth dynamics, geodynamic series*, (Vol. 24, pp. 147–174). AGU.

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

Teixeira Encarnação, J., Visser, P., Arnold, D., Bezdek, A., Doornbos, E., Ellmer, M., et al. (2020). Description of the multi-approach gravity field models from Swarm GPS data. *Earth System Science Data*. https://doi.org/10.5194/essd-2019-158

Thompson, P. F., Bettadpur, S. V., & Tapley, B. D. (2004). Impact of short period, non-tidal, temporal mass variability on GRACE gravity estimates. *Geophysical Research Letters*, *31*(6). https://doi.org/10.1029/2003GL019285

Wahr, J. M., Molenaar, M., & Bryan, F. (1998). Time variability of the Earth's gravity field: Hydrological and oceanic effects and their possible detection using GRACE. *Journal of Geophysical Research*, *103*(B12), 30205–30229. https://doi.org/10.1029/98JB02844

Wiese, D. N., Folkner, W. M., & Nerem, R. S. (2009). Alternative mission architectures for a gravity recovery satellite mission. *Journal of Geodesy*, *83*(6), 569–581. https://doi.org/10.1007/s00190-008-0274-1

Wiese, D. N., Nerem, R. S., & Lemoine, F. G. (2012). Design considerations for a dedicated gravity recovery satellite mission consisting of two pairs of satellites. *Journal of Geodesy*, *86*(2), 81–98. https://doi.org/10.1007/s00190-011-0493-8

Wolff, M. (1969). Direct measurements of the Earth's gravitational potential using a satellite pair. *Journal of Geophysical Research*, *74*(22), 529–5300. https://doi.org/10.1029/JB074i022p05295

Zhong, L., Sosnica, K., Weigelt, M., Liu, B., & Zou, X. (2021). Time-variable gravity field from the combination of HLSST and SLR. *Remote Sensing*, *13*(17), 3491. https://doi.org/10.3390/rs13173491

References From the Supporting Information

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

Ellmer, M. (2018). *Contributions to GRACE gravity field recovery: Improvements in dynamic orbit integration stochastic modelling of the antenna offset correction, and co-estimation of satellite orientations, PhD dissertation*. Graz University of Technology. https://doi.org/10.3217/978-3-85125-646-8d

Gonçalvès, J., Petersen, J., Deschamps, P., Hamelin, B., & Baba-Sy, O. (2013). Quantifying the modern recharge of the "fossil" Sahara aquifers. *Geophysical Research Letters*, *40*(11), 2673–2678. https://doi.org/10.1002/grl.50478

Henry, C. M., Allen, D. M., & Huang, J. (2011). Groundwater storage variability and annual recharge using well-hydrograph and GRACE satellite data. *Hydrogeology Journal*, *19*(4), 741–755. https://doi.org/10.1007/s10040-011-0724-3

Li, H., Reubelt, T., Antoni, M., & Sneeuw, N. (2016). Gravity field error analysis for pendulum formations by a semi-analytical approach. *Journal of Geodesy*, *91*(3), 233–251. https://doi.org/10.1007/s00190-016-0958-x

Mohamed, A., Sultan, M., Ahmed, M., Yan, E., & Ahmed, E. (2017). Aquifer recharge, depletion, and connectivity: Inferences from GRACE, land surface models, and geochemical and geophysical data. *GSA Bulletin*, *129*(5–6), 534–546. https://doi.org/10.1130/B31460.1

Ray, R. D., & Ponte, R. M. (2003). Barometric tides from ECMWF operational analyses. *Annales Geophysicae*, *21*(8), 1897–1910. https://doi.org/10.5194/angeo-21-1897-2003

Rodell, M., Beaudoing, H. K., L'Ecuyer, T. S., Olson, W. S., Famiglietti, J. S., Houser, P. R., et al. (2015). The observed state of the water cycle in the early twenty-first century. *Journal of Climate*, *28*(21), 8289–8318. https://doi.org/10.1175/JCLI-D-14-00555.1

Thomas, A. C., Reager, J. T., Famiglietti, J. S., & Rodell, M. (2014). A GRACE-based water storage deficit approach for hydrological drought characterization. *Geophysical Research Letters*, *41*(5), 1537–1545. https://doi.org/10.1002/2014GL059323

Wahr, J. M. (1987). The Earth's C21 and S21 gravity coefficients and the rotation of the core. *Geophysical Journal International*, *88*(1), 265–276. https://doi.org/10.1111/j.1365-246X.1987.tb01379.x

Watkins, M. M., Wiese, D. N., Yuan, D.-N., Boening, C., & Landerer, F. W. (2015). Improved methods for observing Earth's time variable mass distribution with GRACE using spherical cap mascons. *Journal of Geophysical Research: Solid Earth*, *120*(4), 2648–2671. https://doi.org/10.1002/2014JB011547

Zhao, M., Velicogna, I., & Kimball, J. S. (2017). Satellite observations of regional drought severity in the continental United States using GRACE-based terrestrial water storage changes. *Journal of Climate*, *30*(16), 6297–6308. https://doi.org/10.1175/JCLI-D-16-0458.1