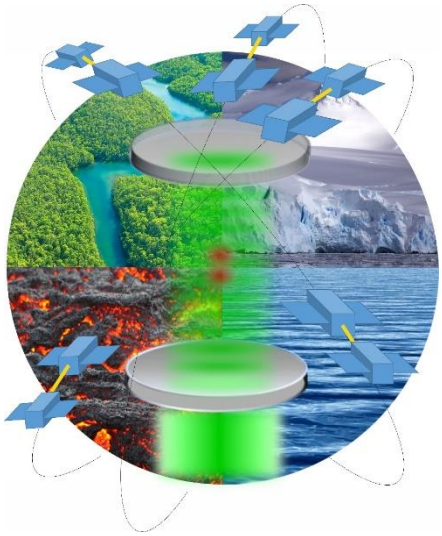
ESA RFP/3-17687/22/NL/SD

Quantum Space Gravimetry for monitoring Earth's Mass Transport Processes (QSG4EMT)

Executive Summary

Gravity field observations are a unique measurement technique to observe and monitor mass transport in the Earth's system. Sustained gravity field observation from space contributes significantly to numerous of Essential Climate Variables (ECVs) as defined by Global Climate Observing System (GCOS), and directly measures changes of the recently adopted ECV "Terrestrial water storage".



Next-generation gravity missions are expected to enhance our knowledge of mass transport processes in the Earth system, establishing their products applicable to new scientific fields and serving societal needs. Compared to the current situation (GRACE Follow-On), a significant step forward to increase spatial and temporal resolution can be achieved by new mission concepts such as the joint NASA/ESA mission concept Mass change And Geosciences International Constellation (MAGIC).

Quantum gravity mission constellations are a core component of ESA's Accelerator "Space for a Green Future". The main advantage of quantum instruments such as cold atom interferometers (CAI) is their close to flat error spectrum, thus reducing high-amplitude long-wavelength errors of classical electrostatic accelerometers.

In the frame of the ESA-funded project "Quantum Space Gravimetry for monitoring Earth's Mass Transport Processes (QSG4EMT)", we analyzed the potential of quantum satellite gravity (QSG) missions to improve the monitoring of mass distribution and mass change processes in the Earth system. The error budget of current gravity satellite missions was analyzed, in order to identify the biggest error contributors, and to evaluate potential future improvements regarding instrumentation. Since instrument errors are not the dominant error source, but rather temporal aliasing errors resulting from an under-sampling of high-frequency temporal gravity signals, extended satellite constellations and improved processing techniques have been investigated with the focus on low-low inter-satellite tracking mission constellations, quantum gradiometry and hybrid concepts, with the goal to reduce temporal aliasing errors. In parallel, user requirements have been formulated. The impact of extended constellations in combination with improved sensor technologies have been assessed for the main applications fields, continental hydrology, climate modelling, oceans and solid Earth.

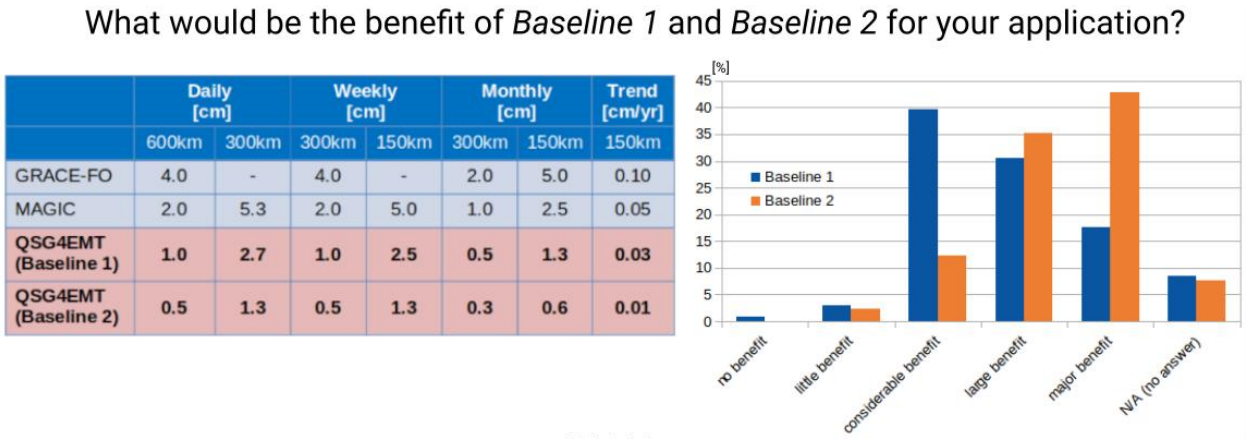
1) Science and user requirements for QSG missions

The goal of this work package (WP100) was the definition of user requirements for future constellations of quantum gravity missions resulting in (1) a list of (new) application fields that can be investigated with such a mission, and (2) performance numbers that would be required from the users' perspective, both leading to an update of the Science and Traceability Matrix (SATM) table. To come up with a consolidated view on user needs and application-dependent science requirements and to identify new users and application fields, we targeted at involving the community as broadly as

possible. Therefore, a community online questionnaire was designed and broadly advertised among various potential user groups. Questions targeted the background of the person, their specific demands regarding the spatial & temporal resolution, latency, and accuracy of mass change products, as well as various topics, e.g. regarding data combination or the assimilation into models. In total, 131 users from 25 different countries spanning all applications fields (hydrology, oceanography, glaciology and ice sheets, solid Earth studies, geodesy, atmosphere & climate modeling) participated.

For the identification of new applications, we added a question to the user questionnaire, that explicitly asked the participants to identify applications that have not been possible with current missions but would be possible with the hypothetical quantum mission scenarios. Answers have been synthesized and new application fields have been extracted from the free-text answers received upon this question and incorporated into the SATM table. When asked about their application driven demands independent of specific mission scenarios („wish list"), users asked for a spatial resolution of 10-100km (40%) to 100-300km (45%) as threshold and <10km (46%) or 10-100km (47%) as desired outcome. For the temporal resolution a peak can be identified for one month (44%), but a considerable number (24%) also prefers one week. For the desired temporal resolution, the majority is interested in a one-day resolution (35%), followed by a weekly resolution (26%). The latency requirements mostly range between one week and three months (threshold) and one day and one month (desired).

While the wish list numbers are partly too optimistic and out of reach for satellite gravimetry, there is still added benefit for the specific application fields, even when the demands cannot be fully satisfied. Therefore, two hypothetical baseline scenarios were defined for a potential future quantum gravity mission with "Baseline 1" referring to a conservative accuracy assumption and "Baseline 2" denoting an optimistic scenario. The table in Figure 1 (left) puts the respective (theoretical) performance numbers in perspective to currently achievable accuracies of the GRACE-FO mission and envisaged MAGIC uncertainties. The right side of Figure 1 summarizes the assumed benefit of the two baseline scenarios for the applications under question in the survey. Already the less ambitious Baseline 1 is considered to be of at least considerable benefit (40%), large benefit (31%) or major benefit (18%). For the more optimistic Baseline 2 scenario, the largest number of participants (43%) expects a major benefit from such a potential new mission. As this acceptance of the hypothetical performance numbers is very consistent across all application fields, we adopted these numbers as the default values in the SATM table as threshold (Baseline 1) and target (Baseline 2) requirements. It was agreed upon with ESA that these numbers shall only be changed for individual applications where there is strong evidence, e.g., from data bases such as the observation requirements provided by the GCOS for their Essential Climate Variables (ECVs) and OSCAR (Observing Systems Capability Analysis and Review Tool) or from the scientific literature.



	Daily [cm]		Weekly [cm]		Monthly [cm]		Trend [cm/yr]
	600km	300km	300km	150km	300km	150km	150km
GRACE-FO	4.0	-	4.0	-	2.0	5.0	0.10
MAGIC	2.0	5.3	2.0	5.0	1.0	2.5	0.05
QSG4EMT (Baseline 1)	**1.0**	**2.7**	**1.0**	**2.5**	**0.5**	**1.3**	**0.03**
QSG4EMT (Baseline 2)	**0.5**	**1.3**	**0.5**	**1.3**	**0.3**	**0.6**	**0.01**

Figure 1: Accuracies of hypothetical mission scenarios (left) and their benefit for applications as stated by participants of the user questionnaire (right).

2) Sensitivity analysis of scientific instrument performance

In this work package (WP300), two measurement concepts have been considered: low-low satellite-to-satellite tracking (ll-SST), and gravity gradiometry. The most beneficial configurations for temporal gravimetry have been analyzed when using electrostatic and quantum instrumentation for both measurement concepts, except for the combination of gravity gradiometry and electrostatic instruments, since this has been demonstrated unfeasible by GOCE. We assumed an analytical model for the error amplitude of quantum accelerometers that is a function of the laser wavelength, number of atoms, interferometer contrast, degree of entanglement, momentum space separation, interrogation time and measurement cycle period. We consider two operational modes for the CAI (cold atom interferometer) accelerometer/gradiometer: *concurrent* atom cloud preparation and interrogation, where the interferometry takes place at the same time as the Bose-Einstein condensate is being prepared, and *sequential* atom cloud preparation and interrogation, where the processes for cloud preparation and interrogation do not overlap, leading to a more extended measurement cycle period. We derived analytical formulas for the amplitude of the Coriolis and centrifugal accelerations as a function of the thermal velocity of the atom cloud, the initial cloud velocity, the angular velocity of the satellite, and attitude compensation provided by steering mirrors.

One additional error source considered in this study is related to the transformation to the Earth-centered, Earth-fixed (ECEF) reference frame. It requires satellite attitude data and is, therefore, not free of errors. Essentially, this error represents how the measured signal contaminates different axes when misoriented in inertial space, and we refer to it as *attitude uncertainty*.

We show that the advantage of additional measurements provided by the *concurrent* operational mode is insignificant compared to the severe reduction of the amplitude of the errors in reconstructing the Coriolis accelerations (henceforth referred to as *Coriolis errors*) given by the *sequential* operational mode. This is because cloud movement is very limited in the latter. In this case, the state-of-the-art attitude measurement systems (here assumed to be the Astrix 200 laser gyroscope combined with a Star Tracker with similar performance to Swarm) combined with an attitude compensation system allow the full accuracy of the CAI instrument to be exploited. This is shown in Table 1, where a simple error propagation considering reasonable values for the input parameters shows no improvement in

Table-1: Standard deviation of the Coriolis term $\boldsymbol{\sigma_{\mathrm{Cor},i}}$, assuming $\boldsymbol{\sigma_{\omega,j} = \sigma_{\omega,j} = 5 \times 10^{-8}}$ rad/s, $\boldsymbol{v}_{\mathrm{atom,therm}} = \mathbf{10^{-6}\ m/s}$ and $\boldsymbol{\sigma_{v_{\mathrm{cloud,initial}}}} = \mathbf{10^{-7} m/s}$ for several combinations of angular velocity compensation scenarios and operational modes (affecting the cloud velocity), for the case of along-track ll-SST and the *i*-axis aligned with the along-track direction.

Attitude compensation scenario	Residual angular velocity [rad/s]	Concurrent mode $[\mathrm{m/s^2}]$	Sequential mode $[\mathrm{m/s^2}]$
		$v_{\mathrm{cloud,k}} = \sigma_{v_{\mathrm{cloud,therm}}} = 2.3\mathrm{nm/s}$ $v_{\mathrm{cloud,j}} = 2.5\mathrm{cm/s}$	$v_{\mathrm{cloud,k}} = v_{\mathrm{cloud,j}} = \sigma_{v_{\mathrm{cloud,therm}}} = 2.3\mathrm{nm/s}$
No tilting mirror	$\omega_{\mathrm{j}} = 1.1 \times 10^{-3}$ $\omega_{\mathrm{k}} = 10^{-4}$	2.5×10^{-9}	2.2×10^{-10}
Minimum pitch-rate compensation	$\delta\omega_{\mathrm{j}} = \omega_{\mathrm{k}} = 10^{-4}$	2.5×10^{-9}	2.8×10^{-11}
Full attitude compensation	$\delta\omega_{\mathrm{j}} = \delta\omega_{\mathrm{k}} = = 5 \times 10^{-7}$	2.5×10^{-9}	2.0×10^{-13}

the Coriolis errors in the *concurrent* mode (second column), even when using attitude compensation (last row).

The attitude compensation system limits the sequential mode, because the Coriolis error is at the level of 10^{-13} m/s^2 as opposed to 10^{-9} m/s^2 in the *concurrent* model.

For quantum accelerometry in ll-SST, we considered CAI parameters that resulted in instrument accuracy of 10^{-13} m/s^2 for the *sequential* mode and $\sim 10^{-14}$ m/s^2 for the *concurrent* mode, as shown in the solid purple and yellow lines, respectively, in Figure 2.

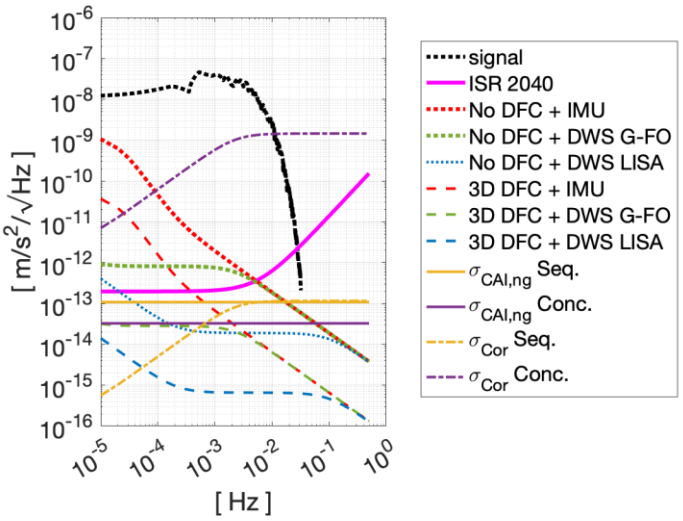


Figure 2. Instrument performance of ll-SST concept for various sensor noise assumptions and error sources

The CAI error amplitudes are both well below the expected inter-satellite error amplitude in 2040, shown in the solid pink line. The Coriolis errors for the sequential and concurrent modes are shown in the dash-dotted purple and yellow lines, respectively. While these errors are not problematic in the case of the *sequential* mode, since they are at most of the same amplitude as the CAI errors, they are destructive for the concurrent mode since they are from 3 to 5 orders of magnitude above. Figure 2 also shows the errors associated with the attitude uncertainty when Drag-Free Control (DFC) is present or not (long or short dashed lines), in combination with attitude derived with a star tracker, laser gyroscope and 3 options for a Differential Wavefront Sensor (DWS), in the red, green and blue lines. If there is no DFC, then DWS at the accuracy of the LISA mission must be used with a slight degradation below 0.03 mHz (short-dashed blue line). If 3D DFC is used, then the DWS currently available in GRACE-FO is sufficient (long-dashed green line).

For CAI gradiometry, considering only pairs of CAI accelerometers operating in the *sequential* model, shown in Figure 3, its sensitivity (red line) is barely enough to resolve time-variable gravity signal up to 3 mHz, corresponding roughly to SH degree 17. The Coriolis errors (yellow line) make it impossible to observe this signal below 0.4 mHz or SH degree 2. The effects of attitude uncertainty (dashed green line) are at least one order of magnitude below the Coriolis effects, and only surpass the magnitude of the gradiometer CAI sensitivity below 0.02 mHz.

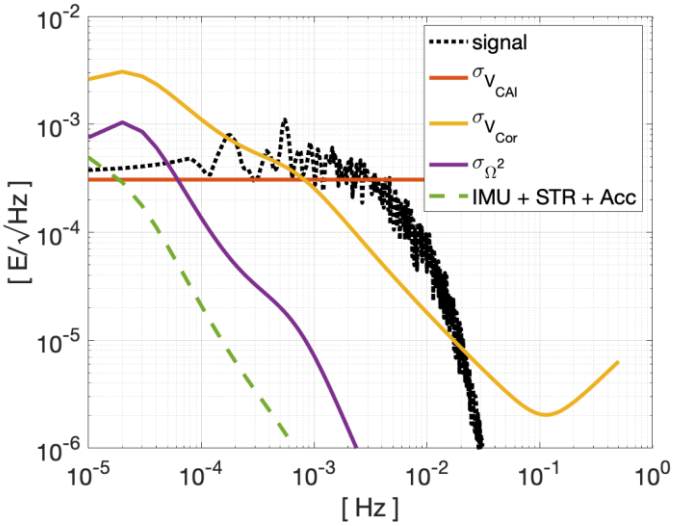


Figure 3. Instrument performance of SGG concept

In contrast to the ll-SST case using quantum accelerometers, where the complete signal spectrum is resolved with a high signal-to-noise ratio, the quantum gradiometer with the same CAI parameters is barely able to determine the time variable signal, with an SNR mostly between 1 and 2, peaking at 3 and dipping at 0.5 at some frequencies. This example reinforces that the high accuracy of all instruments is critical to the success of CAI gradiometry. Although quantum technology may allow for extremely high CAI sensitivities, a proportional improvement of the attitude sensors is necessary.

We demonstrated that the effects of inaccurately measured attitude in the Coriolis accelerations are of paramount importance to the success of the CAI satellite gravimetry. Any CAI concept operating in *concurrent* mode is not limited by its sensitivity, but by the Coriolis effects, even in the quiet environment of space. For demonstration purposes of CAI technology to measure the time-variable gravity field, the best option is the ll-SST measurement concepts and a CAI accelerometer operating in *sequential* mode because this requires less demanding CAI parameters. With the progress of laser metrology, this is still the best option to ensure the accuracy of the LTI instrument is fully exploited since the parallel development of CAI technology allows for comparable accuracies. Of note is the low sampling rate inherently associated with the *sequential* mode, which limits the ability to measure high frequencies. The obvious solution is to consider the hybrid use of CAI and classical accelerometers, which would also allow for validation and calibration of both instruments, of special interest during the initial stages of the exploitation of CAI technology.

3) Trade-space and simulation results of current and extended constellations

The benefit of quantum accelerometers on future satellite gravity missions is assessed based on (1) existing/upcoming satellite gravity mission scenarios (WP300, applying the LL-SST resp. gradiometry principle), (2) on future scenarios incorporating larger constellations (WP400 for LL-SST, WP500 for gradiometry), and (3) based on future mission scenarios applying alternative concepts (WP600). Eventually, the different mission performances and benefits have been assessed for all investigated scenarios (WP300-600). Hence, to achieve a better overview, the executive summary for these WPs is merged to provide a better consistency. This is also in line with the project strategy to create and investigate one single trade-space for all investigated constellations.

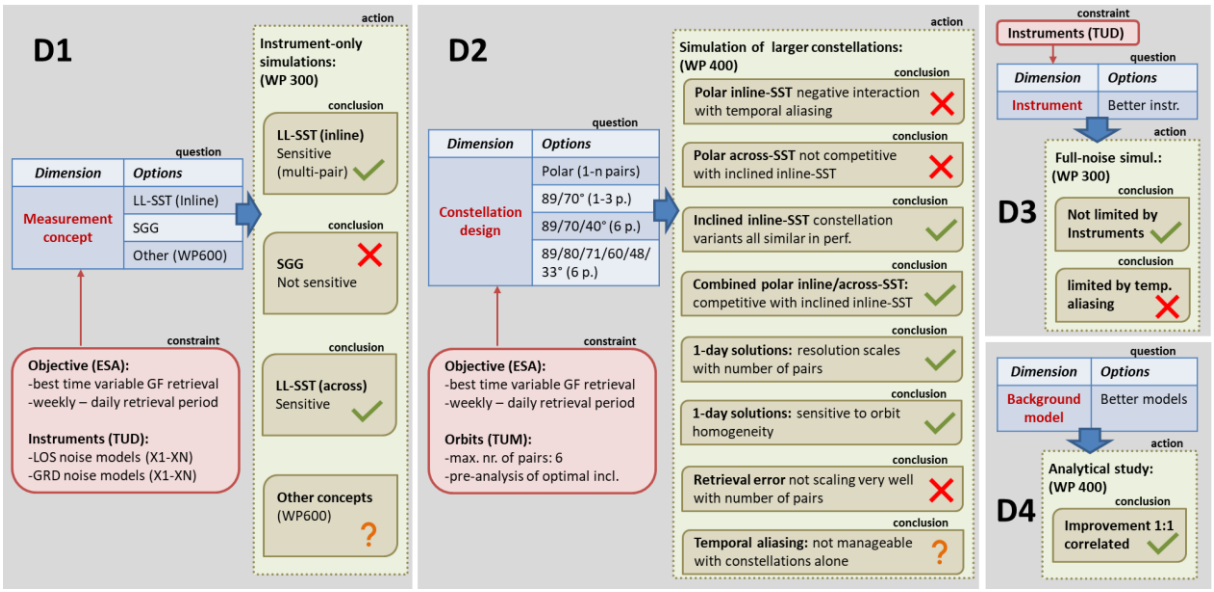
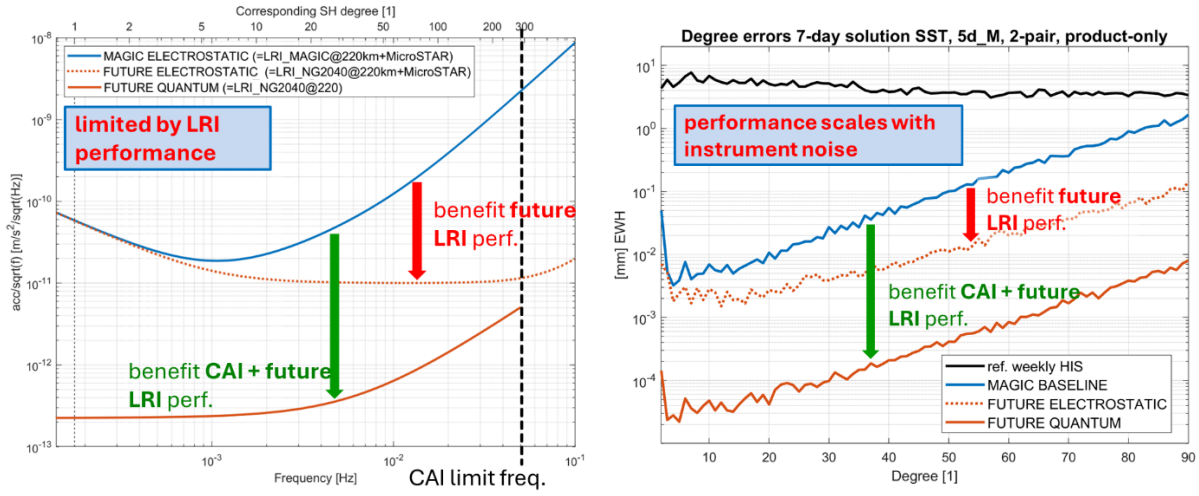


Figure 4 Illustration of the final trade-space with important constraints (red boxes), feasible options (blue boxes) and investigative actions with conclusions (green boxes). Green ticks represent positive conclusions, red crosses negative conclusions (regarding a possible future quantum gravity mission) and yellow question marks still open issues.

The investigated trade-space (see Figure 4) summarizes well the structure of the work performed within WP300-600. It abstractly describes the most important variables (i.e. dimensions) that may impact future satellite gravity missions. Concretely, three main entities (dimensions) have been identified which govern a satellite gravity mission: (D1) the measurement concept (e.g., LL-SST, gradiometry, HL-SST, etc.), (D2) the constellation design (e.g., 1-, 2-, n-satellite/-pair constellations), and (D3) the instrument performances (e.g., accelerometers, ranging instruments). Also, a fourth dimension (D4), the background model performance, has been defined. However, D4 is not seen as a direct mission variable since the background model performance cannot be influenced by the mission itself directly, but has to be provided from external sources (geophysical models). The impact of all dimensions has been investigated by performing a multitude of simulations with different settings (corresponding to points in the trade-space). The results of simulations are validated by retrieving the gravity field through two different approaches.

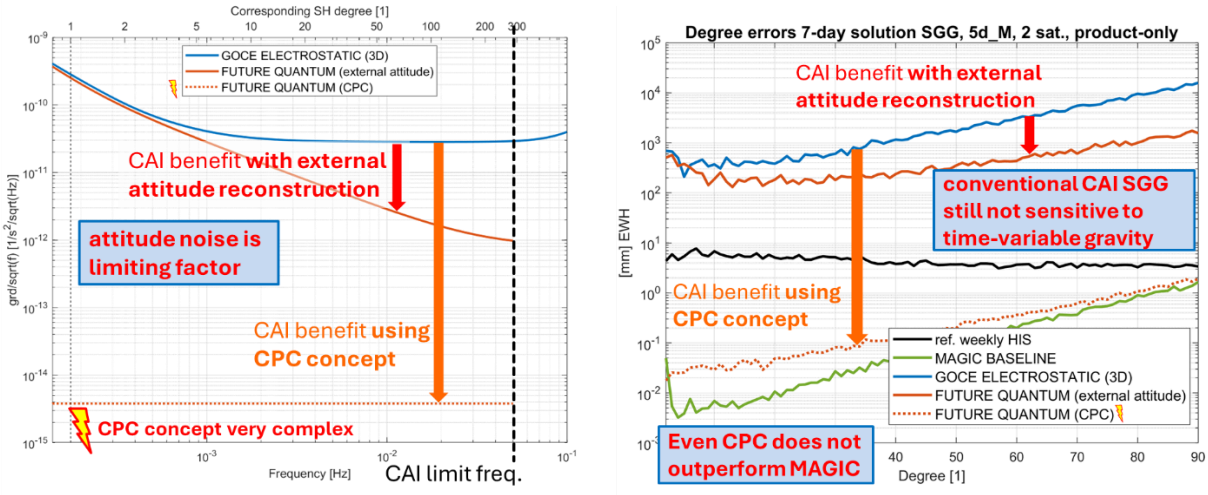
Starting with the measurement concept (D1), the established (inline) LL-SST and SGG principle were mainly investigated (WP300-500). In addition (WP600) also possibly feasible alternatives have been analyzed. Concretely, it has been found that the LL-SST concept is likely the most suited option for future (quantum) gravity missions (see Figure 5). The reason for this is that, compared to SGG, the sensitivity to the gravity field is several orders of magnitude better when deploying the same quantum sensors. On the other hand, SGG requires very high sensitivity from the quantum instrument and a like-wise high accuracy in the reconstruction of the satellite's attitude to become sensitive to Earth's time-variable gravity field (see Figure 6). At the time of writing, the associated technical hurdles seem still too high to consider SGG as a viable alternative to SST (see WP200). As another alternative LL-SST, also high-low SST (HL-SST) has been taken into consideration (in WP600). While HL-SST shares the same sensitivity property than LL-SST, the inter-satellite distances are much larger in comparison. Unfortunately, the performance of all investigated ranging instruments decays proportional with the distance, which is why a future quantum HL-SST mission will likely be governed by the ranging noise and, hence, would not benefit from the high performance of the quantum accelerometers. Hence, for a future quantum mission, there is currently no better option found than LL-SST. However, for LL-SST

there exists theoretically an alternative variant where the satellites of one pair are not orbiting on the same orbital plane following each other (inline SST, I-SST), but on different orbital planes side-by-side (so called across-track SST, A-SST, see Figure 7). The advantage of this concept is the perpendicular measurement direction which could complement the inline measurements from an I-SST mission (to form an IA-SST mission) while orbiting on near polar planes (and, thus, retaining global coverage). It is shown that multi-pair IA-SST missions might pose a real alternative to multi-pair inclined I-SST missions (see Figure 11a). The downside of A-SST is the strongly changing inter-satellite distance and high relative velocities, which might comprise additional technical challenges.



(a) **(b)**

Figure 5 Quantum SST mission performance. (a) Comparison of Amplitude Spectral Densities (ASDs) of the SST observation noise for current-gen. electrostatic (blue) and future quantum accelerometers (orange). (b) Simulated static gravity field retrieval performance in terms of error degree amplitudes for an inclined two-pair mission when assuming the instrument performance from Fig. a.



(a) **(b)**

Figure 6 Quantum SGG mission performance. (a) Comparison of Amplitude Spectral Densities (ASDs) of the SGG observation noise for current-gen. electrostatic (blue) and future quantum accelerometers (orange). (b) Simulated static gravity field retrieval performance in terms of error degree amplitudes for an inclined two-satellite SGG mission when assuming the instrument performance from Fig. a. The Counter-Propagating-Cloud concept is seen as technically too demanding to be a viable option for the near- and mid-term future (see WP200).

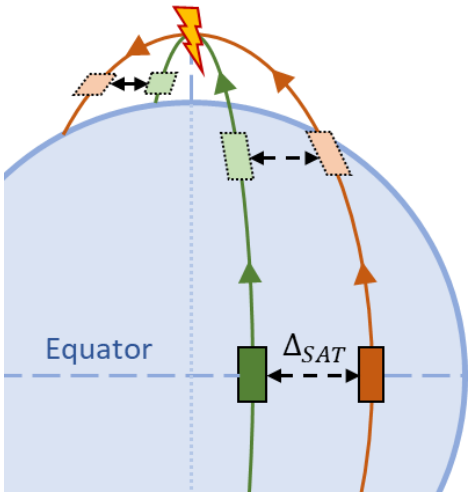
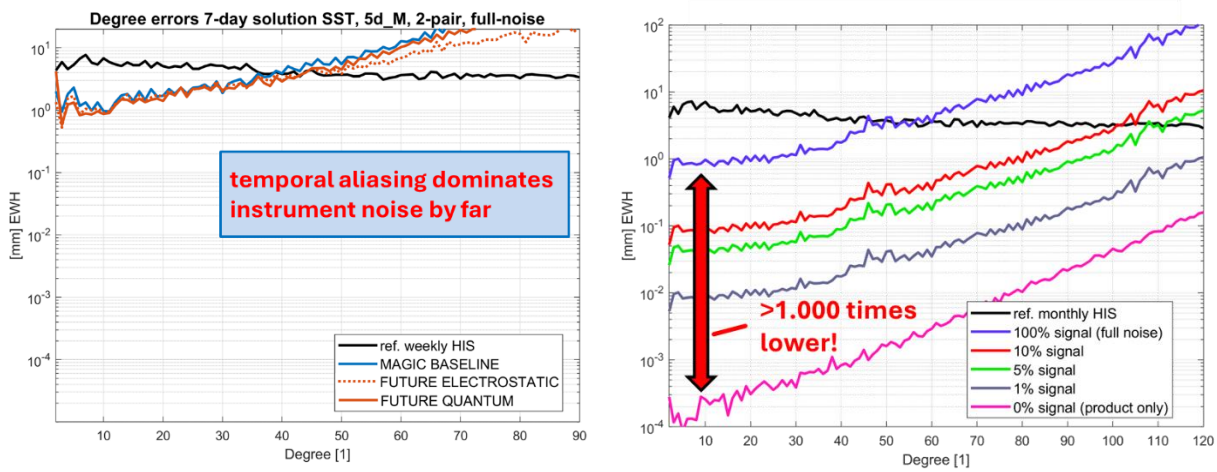


Figure 7 Illustration of the across-track SST (A-SST) concept. The satellite orbits are shifted in the right ascension of the ascending arc (RAAN) to have a certain distance at the equator which reduces towards the poles. At the poles, the intersatellite-distance converges towards zero and the satellites switch side on the descending arc.

It could be observed that better instruments (D3) result in a higher sensitivity to the gravity field signal (cf. Figure 5 and Figure 6). However, under realistic conditions with time-variable gravity signal, all investigated (two-pair/satellite) scenarios are strongly limited by temporal aliasing due to the temporal under-sampling of the gravity field when considering only small constellations (e.g., two pairs, see Figure 8). Hence, it is concluded that the instrument noise is currently not the main limiting factor in the time-variable gravity field retrieval. Only when background model errors (D4) would decrease by a factor of >1000, the full potential of future quantum sensors could be exploited (cf., Figure 8b). Eventually, it is neither expected that background models improve that much nor that future improvements in these models can be significantly influenced by the designed mission (which is why D4 is not a real trade-space variable).



(a) **(b)**

Figure 8 (a) Simulated time-variable gravity field retrieval performance in terms of error degree amplitudes for an inclined two-pair SST mission when assuming the instrument performance from **Figure 5**a. (b) Impact of the background model errors on the time-variable (quantum) gravity field solution (from Fig. a) when assuming n-times better background model knowledge (see legend).

Since temporal aliasing is dominating, larger constellations (D2) need to be considered to increase the spatial-temporal coverage of the mission. In this project, it has been decided to limit the number of satellites/pairs to six. Hence, several different constellations with up to six satellites/pairs have been investigated. Selected constellations for which later on also the impact was assessed are summarized in Table 2, and their ground track distribution is shown in Figure 9.

Table 2: Selected constellations with 7-day near sub-cycle

Acronym	No. pairs	Inclination [°]
IIC2v1	2 (in-line)	89, 70
IIC3v1	3 (in-line)	89, 70, 40
IIC6v1	6 (in-line)	89, 80, 71, 60, 48, 32
PIAC6v2	6 (3 in-line, 3 across-track)	89, 89, 89, 89, 89, 89

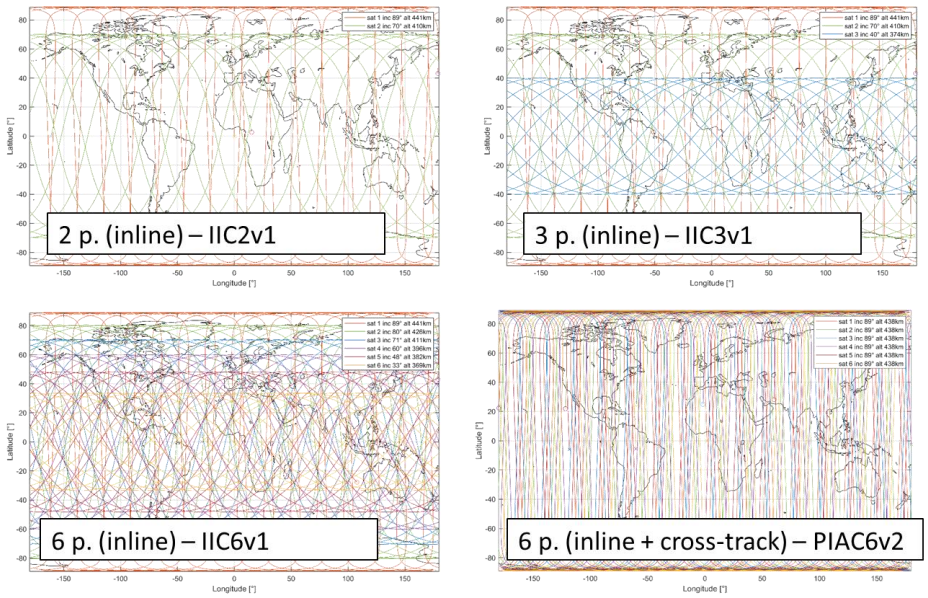


Figure 9 Ground track patterns of selected multi-pair constellations.

The resulting gravity field performance of the 2-pair (IIC2v1), 3-pair (IIC3v1) and 6-pair (IIC6v1) constellation is shown in Figure 10. It is well seen that larger constellations help indeed to mitigate temporal aliasing to some extent (Figure 11a). However, the improvement is far not sufficient to reach the instrument performance level (see Figure 11b). A detailed investigation on the constellation design, the benefits and the limitations can be found in Zingerle et al. (2024).

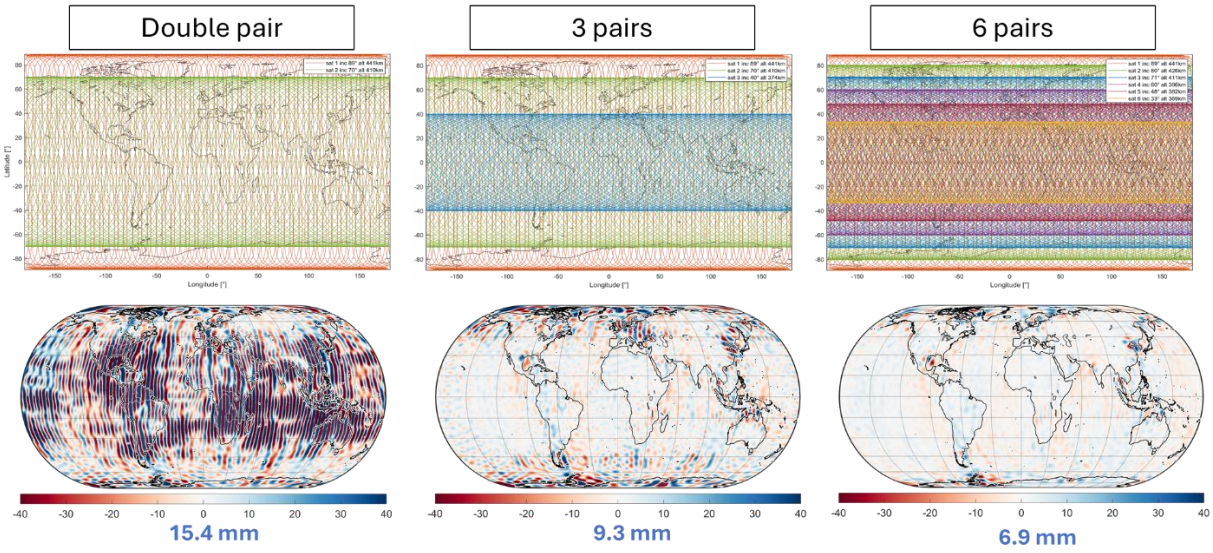
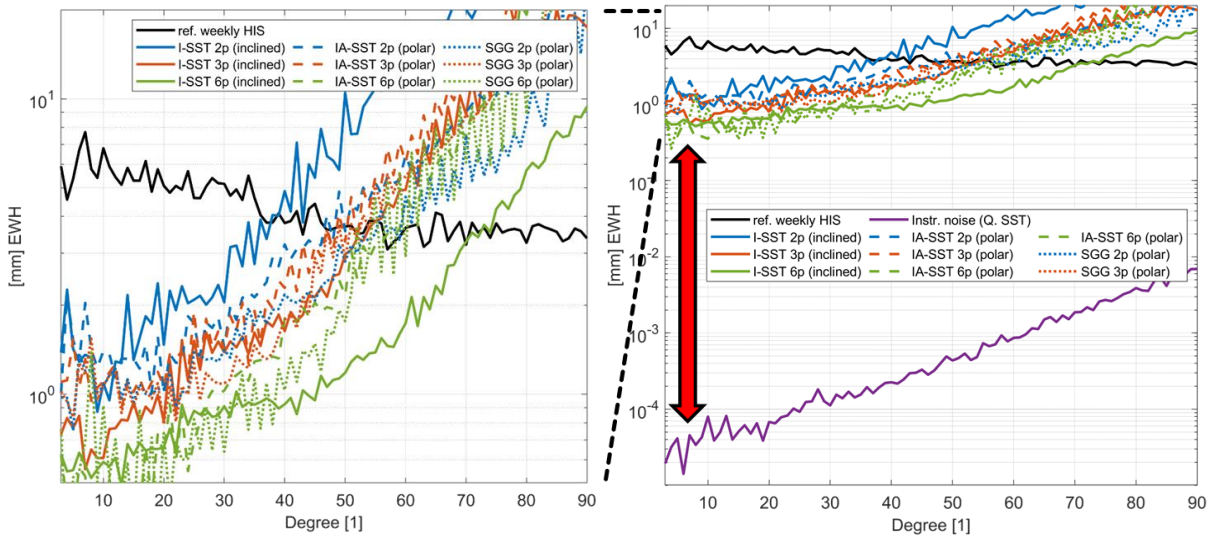


Figure 10 Simulation of larger SST constellations (top row, ground-track after 7 days) and retrieved time-variable gravity field solution (bottom row, 7 day mean) with the global RMS at d/o 60 in terms of equivalent water heights [EWH]. Left: inclined double-pair constellation. Center: inclined triple-pair constellation. Right inclined six-pair constellation.

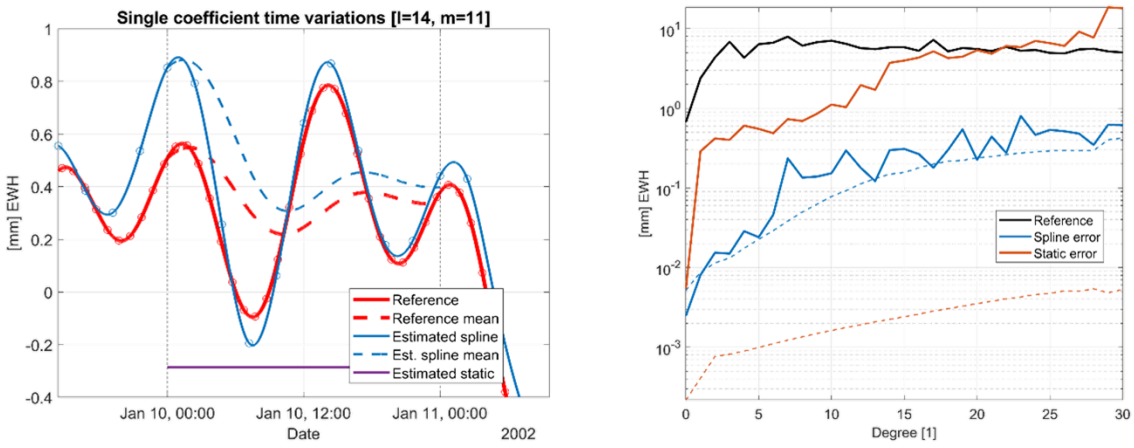


(a) **(b)**

Figure 11 Simulated time-variable gravity field retrieval performance for extended constellations (SGG and SST) in terms of error degree amplitudes of a 7 day solution (see legend). (b) Comparisons of the retrieval errors from Fig. a with the quantum instrument sensitivity (purple line, i.e., static gravity field retrieval error).

This leads to the conclusion that the temporal aliasing problem cannot be solved by larger constellations alone, but that also more sophisticated parameterization schemes are needed in the future to account for the steadily changing gravity signal. In the standard processing, the gravity field is usually only modeled to be static within the retrieval period. So, even a linear trend in the signal cannot be represented correctly and introduces instead an aliasing-like error pattern. To overcome this issue, it is proposed to use a time-aware cubic spline parameterization instead. Splines have been chosen, because they can approximate nearly every function, have local support, are easy to construct and to regularize. An exemplary result is shown in Figure 12. It is seen that the spline approach

performs much better (about one order of magnitude) compared to the static solution. This highlights the importance of not only considering better instruments and larger constellations, but also adapted parameterization schemes. Sufficient benefit, however, can only be achieved with larger constellations to overcome spatial aliasing issues.



(a) **(b)**

Figure 12 Visualization of the spline parameterization on a dedicated 5-pair mission with 1/5th of a day sub-cycle. (a) Time-progression of the reference value of a single coefficient (red) in comparison to the spline solution (blue) and the static solution (violet). Dashed lines depict the cumulative mean. (b) Simulated time-variable gravity field retrieval performance in terms of error degree amplitudes of a daily mean solution when assuming $10^{-12} m/s^2$ white noise instrument behavior. Orange: static solution. Blue: spline solution.

4) Regional solutions

As shown above, satellite constellations for gravity field recovery are promising to provide data allowing for time-variable gravity field investigations with higher accuracy and spatial and temporal resolutions than the current state of the art. In this context, the question arises whether a sequence of global solutions in terms of spherical harmonic coefficients (for example representing Equivalent Water Heights for hydrological applications) is the best option to estimate local signals that may have a stronger amplitude than the global average, and therefore may be inferred with a higher spatial resolution.

The aim of this work package (WP700) was to investigate regional solutions based on collocation by exploiting the space-wise approach. The method basically consists of two steps. Firstly, a global spherical harmonic solution by least-squares adjustment is computed. Then a grid prediction by collocation is performed on the residuals to refine the global solution by exploiting the local characteristics of the gravity field. In this local gridding, modelling signal and noise covariances plays a crucial role. This should be empirically driven by the observations and cannot be done by neglecting the temporal aliasing due to the gravity field variations during the analyzed time span. The method also provides an estimate of the full error covariance matrix of the grid values, which may be useful for subsequent investigations. This matrix is computed by formal error covariance propagation, and it is a-posteriori rescaled based on Monte Carlo simulations.

In order to assess the performance of the space-wise approach in computing regional solutions, some test areas were selected. These are: the East China Sea for ocean applications, some small to large scale river basins (Amazon, Danube, Ganges, Elbe, Rhine, Oder, and Uruguay) for hydrological and

climate applications, and the Bengkulu earthquake that occurred in 2007 for solid Earth applications. For each of these regions, a 7-day time-series of local grids was computed, estimating the TWSA in terms of equivalent water height (EWH) for ocean and hydrological applications and the gravity disturbance at 10 km altitude for geophysical applications.

The solutions were computed considering inline satellite-to-satellite tracking with 2, 3, and 6 pairs of satellites or mixed inline and cross-track satellite-to-satellite tracking with 6 pairs of satellites (for hydrological applications only). Gradiometry was not considered, because its performances are expected to be inferior in a realistic instrumental scenario.

As an example, a comparison between the Amazon total water storage anomaly (TWSA) grid computed from the global least-squares solution and the corresponding local collocation refinement is shown in Figure 13 with 3 pairs of satellites over a 7-day retrieval period. The second step of the space-wise approach leads to an improvement in terms of both accuracy and spatial resolution. To give a more comprehensive picture of the results over the test areas selected for hydrological applications, the average RMSE over a time span of 1 year with 7-day solutions is shown in Figure 14 for the different mission configurations. Further averaging the root mean square error (RMSE) over all regions, it can be stated that the estimation accuracy is always better than 3 cm (in terms of EWH) independently of the chosen orbit configuration. Increasing the number of satellite pairs, the average estimation accuracy can be improved to 2 cm, with some regions reaching 1 cm level.

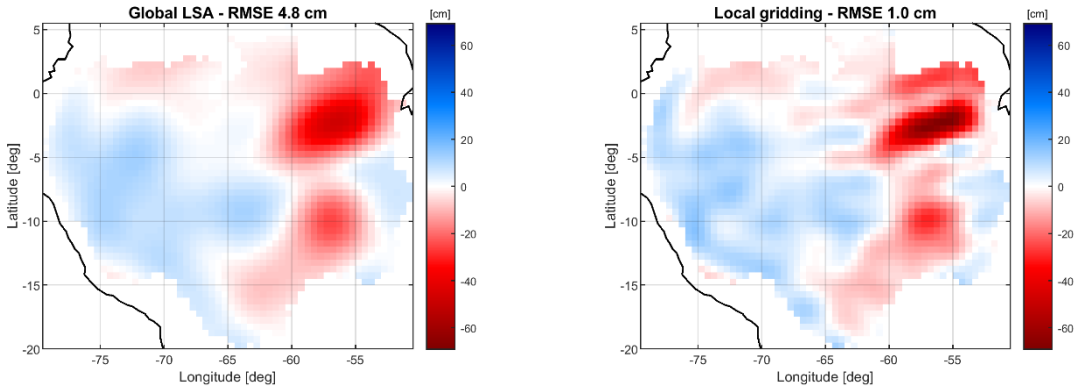


Figure 13 TWSA estimate in terms of EWH for the first week of the year 2002, considering quantum instrumentation with 3-pairs of inline satellites. The Least Squares estimate (regularized, up to d/o 120) is reported in the left panel, while the refined gridding estimate is reported in the right panel.

Regarding Solid Earth Applications, the results of the 7-day solutions computed for the Bengkulu Earthquake showed that the time-variable gravity field can be estimated with a spatial resolution between 1.2° (2-pairs) and 1° (6-pairs), corresponding to a spherical harmonic degree between 150 and 180. As for the estimation accuracy, the RMSE is in the range between 10 μGal and 5 μGal, depending on the orbit configuration. Generally, the higher the number of satellite pairs, the better is the accuracy, even if this may depend on the location of the event being studied.

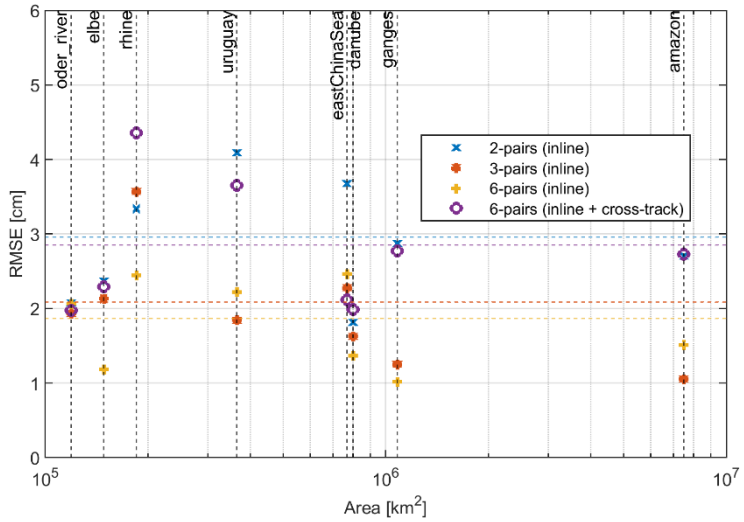


Figure 14 Overall empirical estimation error for each selected region as a function of the basin size, considering regional solutions with a 7-day retrieval period for different satellite configurations. The horizontal dashed lines represent the average of this estimation error over all the regions for the corresponding satellite configuration.

Regarding Solid Earth Applications, the results of the 7-day solutions computed for the Bengkulu Earthquake showed that the time-variable gravity field can be estimated with a spatial resolution between 1.2° (2-pairs) and 1° (6-pairs), corresponding to a spherical harmonic degree between 150 and 180. As for the estimation accuracy, the RMSE is in the range between 10 μGal and 5 μGal, depending on the orbit configuration. Generally, the higher the number of satellite pairs, the better is the accuracy, even if this may depend on the location of the event being studied.

5) Impact assessment

The main objective was to assess detectability of various geophysical signals by the different proposed constellations and payload of the quantum technology instrumentation.

5.1 Solid Earth applications

The activities in this work package (WP800) involved the computation of the Solid Earth geophysical signals, and the assessment of their detectability. The computation of the signals was accomplished for tectonic events including earthquakes, unrests at submarine volcanos, uplift and subsidence of the Alpine range, followed by the identification of geologic bodies and fluid incursion into the porous rocks of sediment basin, and at last the signals associated to deep mantle flows were considered. The Earthquake Signal Repository was formed by the computation of the co- and post-seismic signal from synthetic models of real events. The repository allows to isolate the gravity change due to the earthquake rupture and the subsequent viscoelastic relaxation, between any given timeframe from the source time to the subsequent years. At marine volcanoes, or seamounts, the mass change due to sudden submarine eruptions was computed, for a set of documented events, including the Fani Maoré submarine volcano sudden growth (2014-2015) and the Hunga Tonga-Hunga Ha'apai sudden explosion (2022). Onsite surveys and/or remote sensing, in addition to petrologic analogies to other seamount systems, support the mass change estimates. The vertical movements of the Alps and surrounding regions was computed from a network of GNSS time series, through isolation of the long-term trends in vertical movement, allowing to construct a time-varying surface change model and its gravity effect. Concerning the Lithosphere, using a model of crustal structure in the test area of the Eurasia – Arabia

collision (encompassing the Caucasus and Zagros Mountains and the surroundings basins), the spatial distribution of intra-crustal bodies (e.g. different geologic units, volcanic complexes) and sedimentary basins was isolated. It serves as a target signal for the retrieval of large-scale static structures and to analyze mass changing with different porosity scenarios. For the deep Earth signals, Dr. Bernhard Steinberger (GFZ) provided a model of long-term mantle dynamics and their gravity effect. In this case, the signal to be detected consists in the difference between two model snapshots, 1 Myr apart, re-scaled according to the timeframe of the observation (e.g. the 1-year change in gravity is represented by 1 ppm of the snapshot difference).

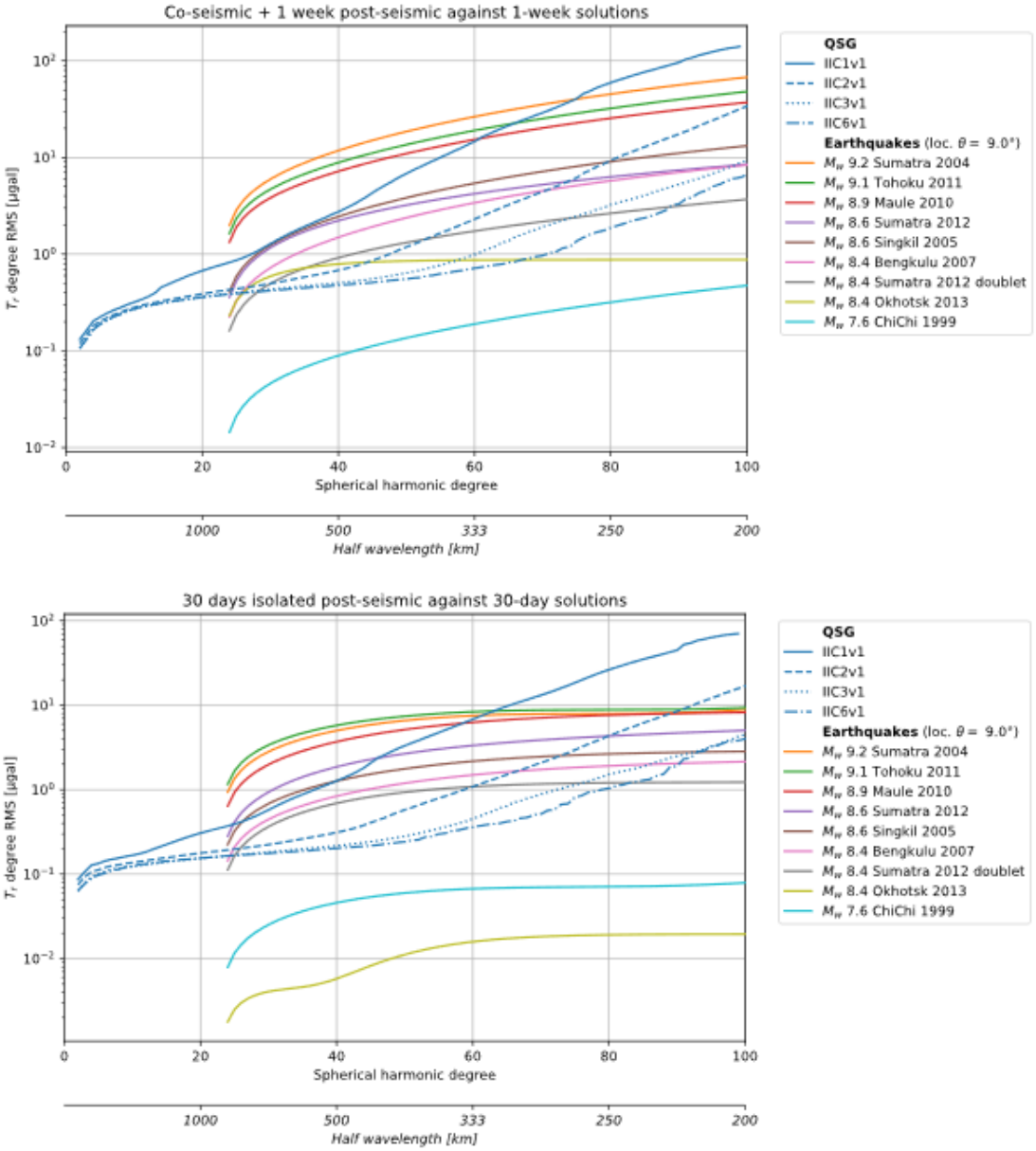


Figure 15 Retrieval error spectra of QSG compared to the localized spectra of the collection of earthquake signals. Short-term detectability of co-seismic and post-seismic change in the first 7 days (top) and 30 days (bottom), respectively after an earthquake, compared with the average retrieval errors. Fields calculated at zero height, cumulative spectrum.

The detectability analyses involved two different strategies: a) comparing the signals (as localized spectra) with the spectral retrieval errors, which are computed as residuals between the observed

(simulated) signal and the average reference signal in the same time interval. In this strategy, which conforms to the type of detectability analysis made in all of the solid Earth signals, the criterion for positive detectability is a SNR > 1 at each spherical harmonic degree at which both the localized signal spectrum and the error degree spectrum are available (see Figure 15).

For the earthquakes we could perform a realistic signal-retrieval analysis of the simulated gravity products, conforming with a workflow that resembles the signal analysis of any real gravity product (e.g., Level-2 global gravity models as available from GRACE). In this strategy (b) the 7-days solutions or averages of more solutions, are compared to the known earthquake signals that are part of the simulations in the updated HIS model. The results show that the two strategies provide consistent results: whenever in strategy a) an earthquake is not visible because the signal spectrum is never above the noise spectrum, in spatial domain (strategy b) the earthquake signal is not above the retrieval error. As an example, the signal of the Maule 2010 is shown in Figure 16.

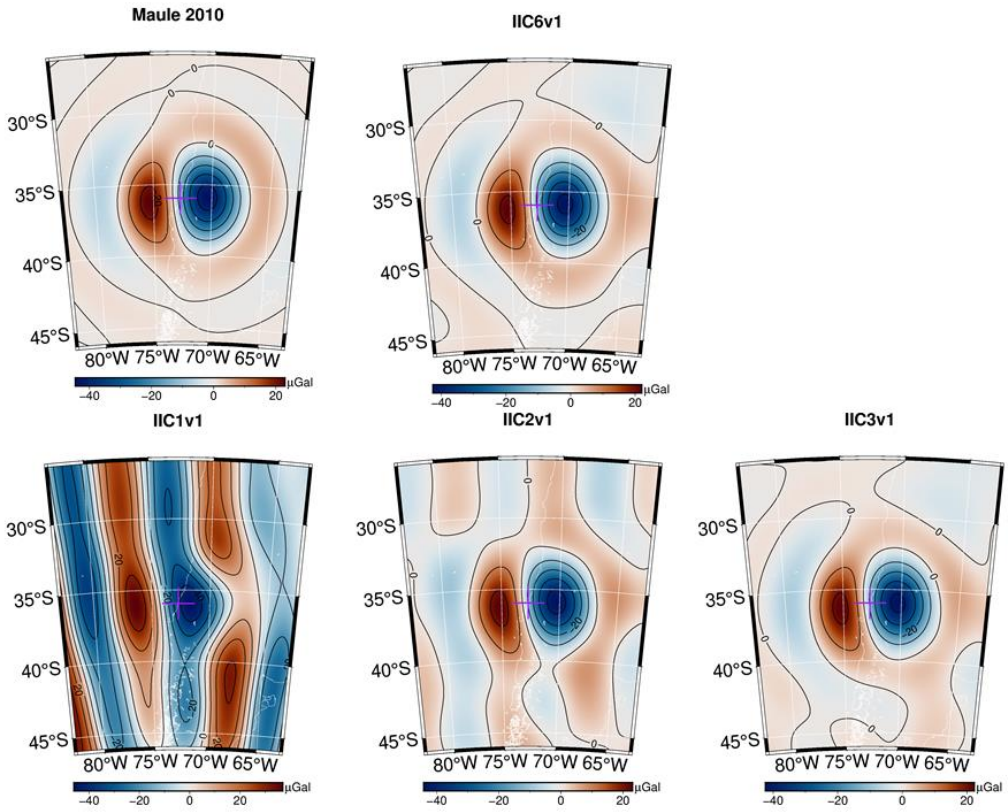


Figure 16 Coseismic gravity change for the Maule 2010 earthquake. The sub-figures are identified by their titles: Maule 2010: synthetic coseismic field. IIC1v1 to IIC6v1: retrieved gravity field covering the time of the earthquake, by taking the difference of the field one week after and one week before the The retrieved signal of the constellation includes the HIS variation.

The detectability of the other geophysical signals is summarized in the following. Seamounts: The spectral analysis predicts that the Hunga Tonga volcano 2022 mass change, at weekly sampling, starts to be visible from the double pair constellation upwards. This is confirmed in the space domain, where the retrieval error for a single pair is about ±140 μGal, reducing to a few units of μGal for the six pairs, which compares to the signal reaching values just above 5 μGal. Alps: The Alps show a gravity change rate smaller than 0.5 μGal/yr, positive over the uplifting range, negative in the Po-basin. We compare

the localized spectrum of the Alpine gravity change in one year, with the noise curves for 1 year, and the same for 4 years. The signal of one year is too small to be detected, it must be accumulated over four years to achieve detectability with a double couple (or better) constellation. Considering the possibility to detect fluid incursion/extraction in a sedimentary basin, be it natural or anthropogenic, we represented it by the spectral signal curves of air filled against water filled porous rock. For the size of the chosen sediment block (5 km deep, 15 000 km^2 area), with 0.1% and 1% porosity, the water mass is 57 Gt and 570 Gt, respectively. The single couple with weekly sampling could not detect the fluid filling of the 0.1% porosity, whereas the double couple could.

The final topic was concerned with the possibility to detect mass movements in the deep Earth, as the mantle flows induced by the history of slab subduction acting on a inhomogeneous mantle The detection of these small signals is challenging, and would be approached with a several decade long time of acquisition (30 years).

5.2 Ocean applications

The imbalance of water mass fluxes into and out of the global ocean leads to a change in total ocean mass, while at regional scale also barotropic mass rearrangements play a significant role. Ocean mass change (OMC) responds to mass loss of the Greenland and Antarctica ice sheets, land glaciers, river discharge driven by terrestrial hydrology, and changes in the land-ocean precipitation–evapotranspiration balance. In combination with measured or modelled sea level variations, estimates of OMC can be utilized to construct sea level budgets, i.e. partitioning total sea level measured by satellite altimetry, to individual mass and steric contributors in order to better understand contemporary and future sea level drivers. This is in particular relevant for steric sea level changes which are related to ocean warming and salinity change, and which are notoriously difficult to obtain from in-situ measurements for the deep ocean. Approaches exist for retrieving contributions either individually and/or as a residual or jointly in an inverse approach. In WP900, we utilize the inverse approach as it involves an explicit weighting between gravimetric spherical harmonic coefficients and (along-track) altimetric sea surface heights based on a-priori error information, and thus, facilitates error propagation for scenarios. The U Bonn approach works via least-squares fitting of several hundreds of dominating (a-priori) spatial fingerprints of sea level change to monthly gravimetric and altimetric data; these fingerprints include static 'passive' sea level patterns related to ice sheet, glaciers and land hydrological changes derived through the sea level equation. Furthermore, steric patterns that we derive from an EOF decomposition of modelled thermo- and halosteric expansion, and patterns that we interpret as internal mass variability and that are mostly related to barotropic motions in the major ocean basins. We replace the GRACE/-FO error model in our real-data inversion processing by the normal equations from each of the four scenarios introduced above, while keeping the radar altimeter error model (assuming Jason-1 and -2, with errors 5-10 cm derived from binning 20Hz data into 1Hz blocks standard deviation) unmodified. While ocean mass errors are already small with the MAGIC mission concept, we find the benefit of future scenarios mainly in their improved ability to separate different contributions to the sea level budget. Figure 17 shows a map of error level improvement of the ocean mass change component of the 3- and 6-pair quantum simulation scenarios relative to the MAGIC-like simulation scenario.

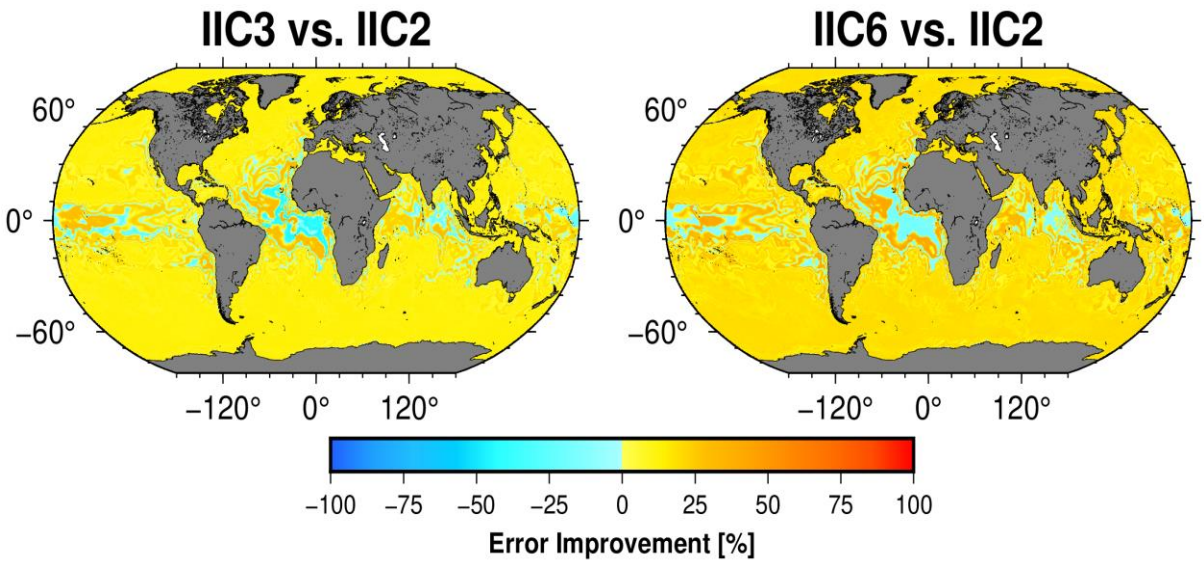


Figure 17 Map of error level improvement of the ocean mass change component derived from 3- and 6- pair quantum simulation scenarios IIC3 and IIC6 relative to the MAGIC-like scenario IIC2.

While the global sea level budget is often considered as closed, it is unclear whether this holds true for regional sea level budgets. These budgets are generally more difficult to derive, due to a variety of local physical effects, as, e.g., wind-driven sea level changes or sedimentation, but also retrieval challenges related to satellite gravimetry and altimetry. Earlier studies pointed out problems and demonstrated limited agreement in many areas. The improved spatial resolution associated with future gravity missions is expected to aid in better closing budgets at regional scale. Therefore, we consider also the common "direct" method as this enables to study the impact of spatial resolution for challenging regions in a more straightforward way. To derive regional OMC from the simulated (i.e. ESAESM) gravity fields, we add the GIA correction to the spherical harmonic coefficients, apply DDK3 filtering and restore the de-aliasing product. After converting to water heights and computing the basin average, we apply a leakage correction derived from DDK3-filtering at the LSDM ESAESM hydrology. Basin averages were studied for the East China Sea (ECS), a Western Pacific marginal sea with an area of 770,000 km². What makes the ECS unique and challenging for budget studies is its complex ocean current system, mostly shallow bathymetry and a large amount of sediments transported by rivers or resulting from coastal erosion. In-situ observations, as, e.g., tide gauge data, for external validation are relatively sparse. Simulation results are compared to the truth from the ocean component of the ESAESM; we find consistent improvements with advanced mission concepts as compared to the GRACE/-FO mission. : RMSE values reach from 1.8 cm (GRACE-FO-like scenario) to 1.5 cm (MAGIC-like scenario) and 1.3 cm for IIC3v1 and IIC6v1.

5.3 Short-term hydrology

Data assimilation (DA) provides a way to integrate satellite-derived total TWSA with hydrological model simulations in a statistically optimal manner. It allows to correct model and forcing data deficiencies at scales where gravimetry provides credible information, while preserving high-resolution information from modelling. DA also enables one to disaggregate TWSA into compartmental storages, and integrate remote sensing data such as snow cover maps in model simulations. In this WP, we investigated the potential benefit of the MAGIC and quantum scenarios compared to the GRACE/-FO mission in an Ensemble Kalman filter (EnKF) assimilation framework, with and without modifying the current DA setup that was developed for integrating GRACE/-FO observations with the WaterGAP model. Different from GRACE/-FO, we find that for MAGIC and quantum scenarios we can assimilate

TWSA maps indeed at the native 0,5° resolution of the model, due to improved resolution of these missions over GRACE/-FO. We thus assimilate monthly TWSA simulated with WaterGAP and perturbed with errors that correspond to the four scenarios as discussed above. We implement the DA via the Parallel Data Assimilation Framework (PDAF) framework, which had been coupled with WaterGAP in the online mode for numerical efficiency. Data uncertainty representation in the DA is simplified via a single scaled SHC normal matrix, propagated to the grid while applying the DDK5 filter, while the uncertainty of the hydrological model (i.e. climate forcing and parameter uncertainties) is kept as in our real-data GRACE/-FO analyses. Due to the expensive computational costs we focus exemplary on South America; but results have been found as transferable to other continents. We find that DA clearly benefits from gridded products from advanced mission scenarios, and this is not limited to spatially downscaling TWSA but also to the retrieval of groundwater, surface water and soil moisture anomalies (see Figure 17). As groundwater is the largest contributor to TWSA, the strongest improvements are found for groundwater as compared to surface water and soil moisture. However, we caution that these results also depend on assumptions of model errors which we kept conservative here.

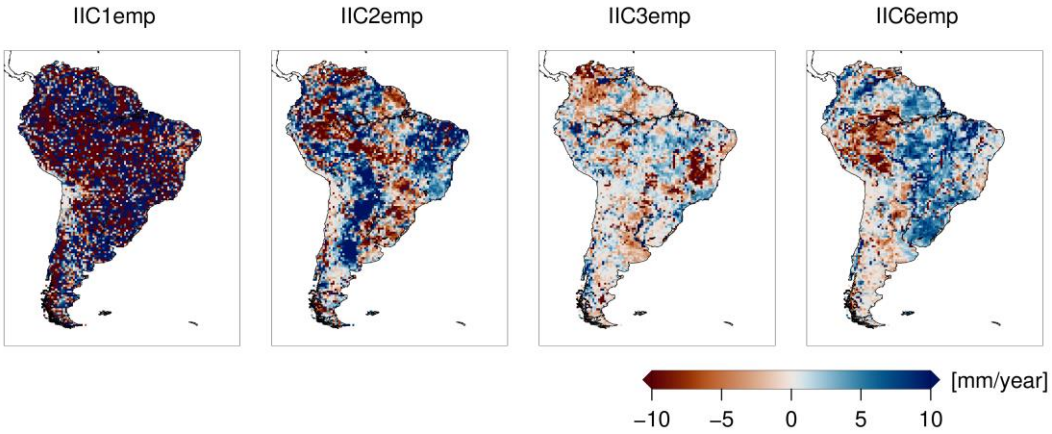


Figure 18: Differences of linear TWSA trends [mm/year] for South America derived from the assimilation of WaterGAP TWSA with formal uncertainties of the IIC1, IIC2, IIC3 and IIC6 scenarios towards the WaterGAP model simulations.

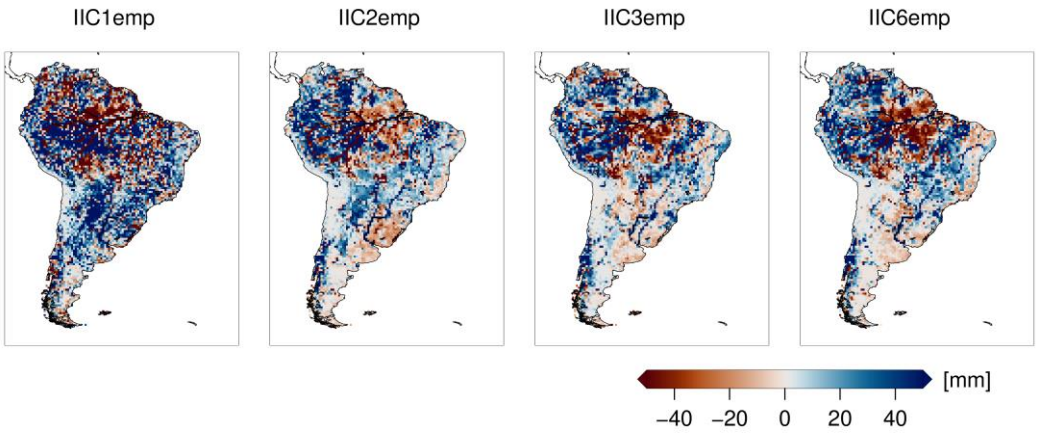


Figure 19: Differences of annual TWSA amplitudes [mm] for South America derived from the assimilation of WaterGAP TWSA with formal uncertainties of the IIC1, IIC2, IIC3 and IIC6 scenarios towards the WaterGAP model simulations.

Land surface models (LSMs) represent the coupled cycling of water, energy, and bio-geophysical matter such as carbon within vegetation and soils. They simulate land-atmosphere interactions and provide thus a key component of climate models. In LSMs, assimilation of satellite-derived TWSA enables one to inform the entire vertical profile of soil moisture, in contrast to in-situ and satellite soil moisture (SM) or land surface temperature (LST) assimilation, and, e.g., to identify problems in soil processes representation. Next to the global study mentioned above, in WP900 we investigate the effect of the scenarios for assimilation with the Community Land Model (CLM). We use the TerrSysMP-PDAF assimilation framework, which allows to integrate observations of LST, SM, groundwater head, and others jointly with TWSA. CLM is set up for the Euro-CORDEX region at 12km grid scale and forced by COSMO-REA6 meteorological data. In our experiments, we use the most recent version CLM version 5.0 (CLM5) as the 'truth' and assimilate into CLM3.5; in other words, we seek to correct for random and non-random deficiencies of the model that is implemented in the DA. Results are compared at (aggregated) grid scale for TWSA, and to basin-averages for the major catchments in Europe with area down to about 50,000 km², and for the monthly storage -- discharge budget which corresponds to the precipitation -- evapotranspiration deficit and has been often used for evaluating atmospheric re-analyses. On catchment scale, the extended quantum gravity constellation scenarios indeed perform better than GRACE/-FO in terms of RMSD and correlation with respect to the synthetic truth. Averaged across all catchments, the RMSD relative to the synthetic truth is 8% lower for scenario IIC6 compared to scenario IIC1. The most significant improvements are observed in Eastern Europe. For instance, in the Narva-Jogi catchment, the RMSD decreases from 39.7 mm for scenario IIC1 to 32.5 mm for scenario IIC6, while the correlation coefficient increases from 0.88 to 0.93, respectively. In some regions over Europe also the representation of trends is improved. In this framework, we also developed literature studies of the potential benefit of improving LSMs through integration of space gravimetric data in the context of coupled modelling, and on their use in operational e.g. forecast services.

5.4 Long-term hydrology and climate

One of the most common hydrological applications of satellite gravimetry is the analysis of the time series of water storage variations in hydrological units such as river basins or aquifers. As an estimation of the achievable accuracy of the mission scenarios, temporal root mean square differences (RMSD) were computed between simulation output and reference time series for basin averages of 405 major world-wide river basins. Figure 19 shows a scatter plot of the RMSD values vs. the sizes of the river basins. Especially the strong improvement of the future constellations over the current GRACE-FO mission becomes evident by the green dots being one order of magnitude higher than the others. While for MAGIC the RMSD of 70% of the river basins is below a threshold of 2.3 cm, the smaller numbers of 1.5 cm (IIC3v1) and 1.0 cm (IIC6v1) show the improvements achievable by the quantum mission constellations, here shown for a spatial resolution of ~220km corresponding to a spherical harmonic truncation degree of N=90.

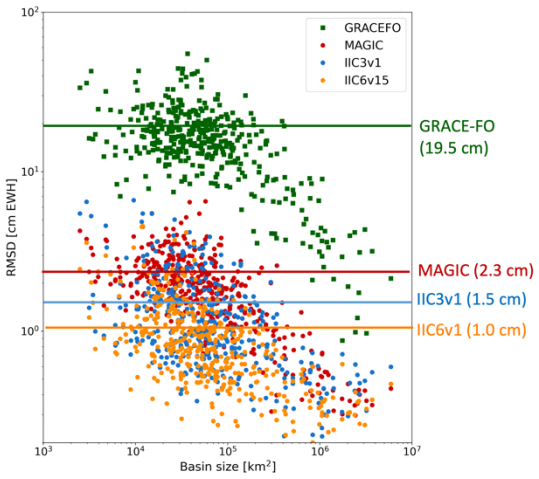


Figure 20: Scatter plot of RMSD of basin average time series vs. size of the river basin for different mission scenarios for DDK5 filtered solutions truncated at N=90 (~220 km), including the filter omission error, i.e. compared to the unfiltered reference up to the same degree. The horizontal lines indicate a threshold for which the RMSD of 70% of the rivers basins is smaller.

To assess the impact of the future mission constellations on climate applications, the performance of long-term trend estimates was assessed from the available 12 years of simulation data. While the maps of post-processed (filtered) linear trend estimates are very similar for all mission scenarios, the benefit of the quantum missions is revealed when increasing the spatial resolution, as these missions are expected to provide a reasonable trend estimation even without applying any filtering. In addition to deriving the trend estimate from a time series of short-term simulations (here: 7 days solutions) the benefit of a direct trend parameterization from the full 12 years of satellite data was investigated and a visible improvement discovered for the two quantum scenarios IIC3v1 and IIC6v1. The presence of inter-annual variations was identified as a major challenge for a robust trend estimate from a comparably small time span. To investigate this issue, a model study was set up based on an ensemble of CMIP6 climate models which revealed a considerable better agreement of short-term trends with long-term (200 years) trend estimates when 50 years (e.g. with 10 years of a quantum mission launched in the 2040s) are available instead of the currently available 20 years of data.

Besides linear trends, also the detectability of climate-related changes to the seasonal cycle are of interest to climate scientists, as they might, e.g., be related to an intensification of the global water cycle revealed by an increase of the annual amplitude. The ensemble of CMIP6 climate models was used to estimate expected changes in the annual amplitude. These expected changes were then challenged against accuracy estimates of measured amplitude changes from 30 years of satellite data. The accuracy estimates were derived by error propagation from empirical accuracies (RMSD of reference vs. simulation time series). In Figure 21 colored pixels denote regions where the projected amplitude change exceeds the magnitude of the accuracy. In this case we assume the amplitude change to be "detectable". It can be seen, that a GRACE-like mission with the chosen weak DDK5 filtering cannot detect the anticipated amplitude changes apart from some very few grid cells. The MAGIC mission, however, already performs much better, with amplitude changes being detectable in 57% of the land area after 30 years of observations for the given setting. The quantum constellations show a considerable added benefit with a detectability in 72% (IIC3v1) and 77% (IIC6v1) of the continental area. Particularly the latter leaves only a few desert areas for which the anticipated amplitude changes are not detectable.

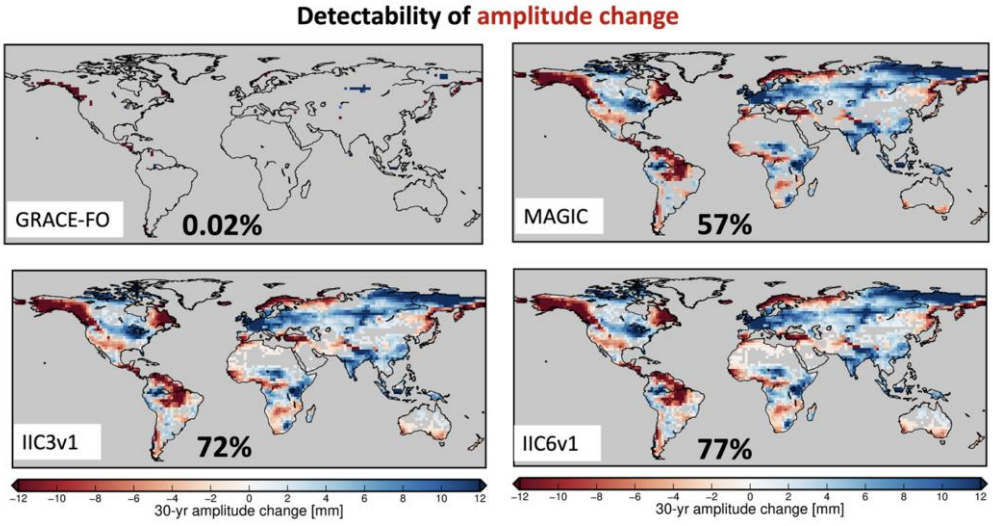


Figure 21: Detectability of amplitude change after 30 years based on projected changes from CMIP6 models and error propagation from mission simulations (7-days solutions, truncated at N=90, DDK5 filtered). Colored pixels: amplitude change is regarded as detectable.

6) Outreach

The main results of this project QSG4EMT are summarized in 3 scientific papers, which are either already published or are going to be submitted soon (status: Sept. 2024):

- Encarnação J., Siemes C., Daras I., Carraz O., Strangfeld A., Zingerle P., Pail R. (202x): Towards a realistic noise modelling of quantum sensors for future satellite gravity missions. In preparation for re-submission to Advances of Space Research.
- Zingerle P., Gruber T., Pail R., Daras I. (2024): Constellation design and performance of future quantum satellite gravity missions. Earth Planets Space 76, 101. https://doi.org/10.1186/s40623-024-02034-3
- Kusche J. et al. (202x): Benefit of multi-pair quantum satellite gravity missions in Earth science applications. In preparation for submission to Nature Reviews Physics.

Additionally, further scientific papers on the results of the user survey, the regional solutions and spline parametrization schemes are planned to be written as an offspring of this study.

7) Conclusions

Instrument concepts for satellite-to-satellite tracking (SST) and gradiometry missions and corresponding performance estimates have been established and corresponding sensitivity analysis performed. Under realistic error assumptions, the benefit of quantum/hybrid sensors is higher for inter-satellite ranging concept than for gradiometry. For SST concepts, the LRI performance has to be further improved to match the quantum/hybrid ACC target performance. For gradiometry, very stringent requirements for the accelerometer and attitude performance apply (accelerometer error level of10^{-15} m/s^2/$\sqrt{\text{Hz}}$, 12 uniaxial gradiometers required for attitude recovery following the counter-propagating cloud concept). Alternatively, the newly developed mission concept of cross-track ranging is an interesting option, but it performs on the same level as in-line constellations.

Compared to the instrument errors, the impact of temporal aliasing is factor of 15 (current instrumentation) to 1000 (QSG instrumentation) higher and thus dominates in all cases the total error budget. Therefore, further improvements in instrument performance have to be complemented by strategies of temporal aliasing reduction, such as extended constellations and improved processing strategies. Therefore, extended constellations of up to 6 pairs have been numerically simulated. They show a gradual performance improvement in terms of temporal aliasing reduction with increasing number of pairs, e.g., a factor of 2.5 from a double-pair to a six-pair constellation, which alone is not sufficient to really make benefit of improved quantum instrumentation.

Improved space-time parametrization strategies such as along-track spline parameterization have been assessed. They show great potential, but they can be applied only for larger constellations to avoid spatial aliasing. Also, a regional gravity retrieval approach has been investigated, which is tailored to regional signal characteristics. It could be demonstrated that regional approaches have the potential to improve global solutions.

In the frame of QSG4EMT, an international user survey with more than 130 participants has been performed. The results of this survey supported the formulation of QSG user requirements. The impact of extended QSG constellations (up to 6 pairs) were evaluated in various applications. In several solid Earth applications, such as big earthquakes, growth of seamounts, tectonic uplift and mantle dynamics, the impact of these different constellations were assessed. The capabilities of extended constellations

were also investigated for ocean applications. An inverse approach to separate sea level drivers was applied. Data assimilation (DA) experiments were performed to integrate satellite-derived total water storage anomalies (TWSA) with hydrological and land surface model simulations in a statistically optimal manner. To assess the impact of the future mission constellations on climate applications, the performance of long-term trend estimates was investigated based on the output of a 12 years numerical simulation. Generalizing the results for all investigated application fields, the scientific return increases when extending the constellation. The greatest impact is the step from a polar single to a Bender double pair constellation, and significant further performance improvements can be achieved by adding a third pair in a rather low inclination of 45-55°. Further increasing the constellation (in our study up to 6 pairs) gradually improves the results further. Also from an application point of view, the impact of extending the constellation is much higher than the inclusion of QSG sensors.

To summarize the main conclusions in one sentence:

Quantum/hybrid sensors are a promising technology for future mass change monitoring from space which needs to be embarked at optimized satellite constellations and complemented by processing strategies to fully exploit their metrology advance in scientific applications.