

An LSC-based framework for geoid-consistent geometric height reference surface determination in the Baltic Sea using SWOT altimetry



Sander Varbla^{1,2} and Thomas Gruber¹
sander.varbla@taltech.ee



¹ Chair of Astronomical and Physical Geodesy, Technical University of Munich, Munich, Germany
² Department of Civil Engineering and Architecture, Tallinn University of Technology, Tallinn, Estonia



INTRODUCTION

- Despite significant theoretical advances in high-resolution gravimetric geoid modelling, establishing a consistent offshore geoid-based height reference surface remains challenging in regions where marine gravity observations are sparse or difficult to obtain.
- High-resolution (2 km × 2 km in this study) [Ka-band Radar Interferometer \(KaRIn\)](#) sea surface height (SSH) observations from the [Surface Water and Ocean Topography \(SWOT\)](#) altimetry mission offer an alternative geometric approach to marine geoid determination.
- Combined with hydrodynamic modelling ([Global Ocean Physics Analysis and Forecast data product](#) in this study), instantaneous geoidal heights can be derived by removing modelled dynamic ocean topography from SWOT SSH observations.
- The ensemble of instantaneous geoidal heights can be used to construct a static, geoid-consistent height reference surface.

METHOD (DISCRETE ESTIMATES)

- The SWOT SSH observations are first corrected by applying solid-Earth and pole-tide corrections, as well as a crossover correction that accounts for cross-track tilts. These corrections are provided by the SWOT data centre.
- The corrected SSH observations are converted into geoidal height estimates by removing hydrodynamic-model-derived instantaneous dynamic ocean topography:

$$N^{inst}(\varphi^{SWOT}, \lambda^{SWOT}, t) = SSH_{corr}(\varphi^{SWOT}, \lambda^{SWOT}, t) - DT_{HDM}(\varphi^{SWOT}, \lambda^{SWOT}, t)$$

- Due to the significant long-wavelength component of the geoid signal, the instantaneous estimates are not directly suitable for covariance modelling or subsequent gridding using least-squares collocation (LSC) principles. A reduction to residual values using a suitable reference field (e.g., using EIGEN-6C4 evaluated up to degree 2190, as in this study) must be performed:

$$r^{inst}(\varphi^{SWOT}, \lambda^{SWOT}, t) = N^{inst}(\varphi^{SWOT}, \lambda^{SWOT}, t) - N_{REF}(\varphi^{SWOT}, \lambda^{SWOT})$$

- The resulting point cloud of instantaneous residual values $r_k^{inst} = r_k^{inst}(\varphi_k^{SWOT}, \lambda_k^{SWOT}, t_k)$ (k indexes individual SWOT observations) can be excessively dense for efficient computations. Spatiotemporal weighted average and standard deviation (superscript *avg* denotes consistency) are thus computed within a regular grid composed of cells C_{ij} (e.g., 2 km × 2 km, as in this study; i and j index grid cells in latitude and longitude directions, respectively), yielding time-independent estimates:

$$r_{ij}^{avg} = \frac{\sum_{k \in C_{ij}} w_k r_k^{inst}}{\sum_{k \in C_{ij}} w_k} \quad \text{and} \quad \sigma_{ij}^{avg} = \sqrt{\frac{\sum_{k \in C_{ij}} w_k (r_k^{inst} - r_{ij}^{avg})^2}{\sum_{k \in C_{ij}} w_k}}$$

- Weights are defined as $w_k = 1 - d_k^{SD}/\sqrt{8}$, where d_k^{SD} is the spatial displacement between the nominal and actual SWOT grid coordinates.
- Instantaneous residual values r_k^{inst} satisfying the condition $|r_k^{inst} - r_{ij}^{avg}| > 3\sigma_{ij}^{avg}$, for $k \in C_{ij}$, are classified as outliers. r_{ij}^{avg} and σ_{ij}^{avg} are recomputed after the removal of outliers.
- An additional leave-one-out-based local outlier detection stage is applied to assess the consistency of the resulting cell-based averaged residual values r_{ij}^{avg} . Outliers are identified using a criterion $|r^{sus}| > |r^{prox}| + 3\sigma[r_{ij}^{avg}]$, where $\sigma[\cdot]$ denotes the standard deviation over estimates r_{ij}^{avg} for $(i, j) \in Q$.
- Q is the set of nearest estimates r_{ij}^{avg} (e.g., a total of 250, as in this study) to a suspected outlier r^{sus} .
- Proxy estimate r^{prox} is computed using inverse distance weighting interpolation.
- Outlier detection is iterated as long as the condition $o_m - o_{m-1} > o_1/10$ holds, where o is the number of identified outliers, and the subscript shows the iteration number.

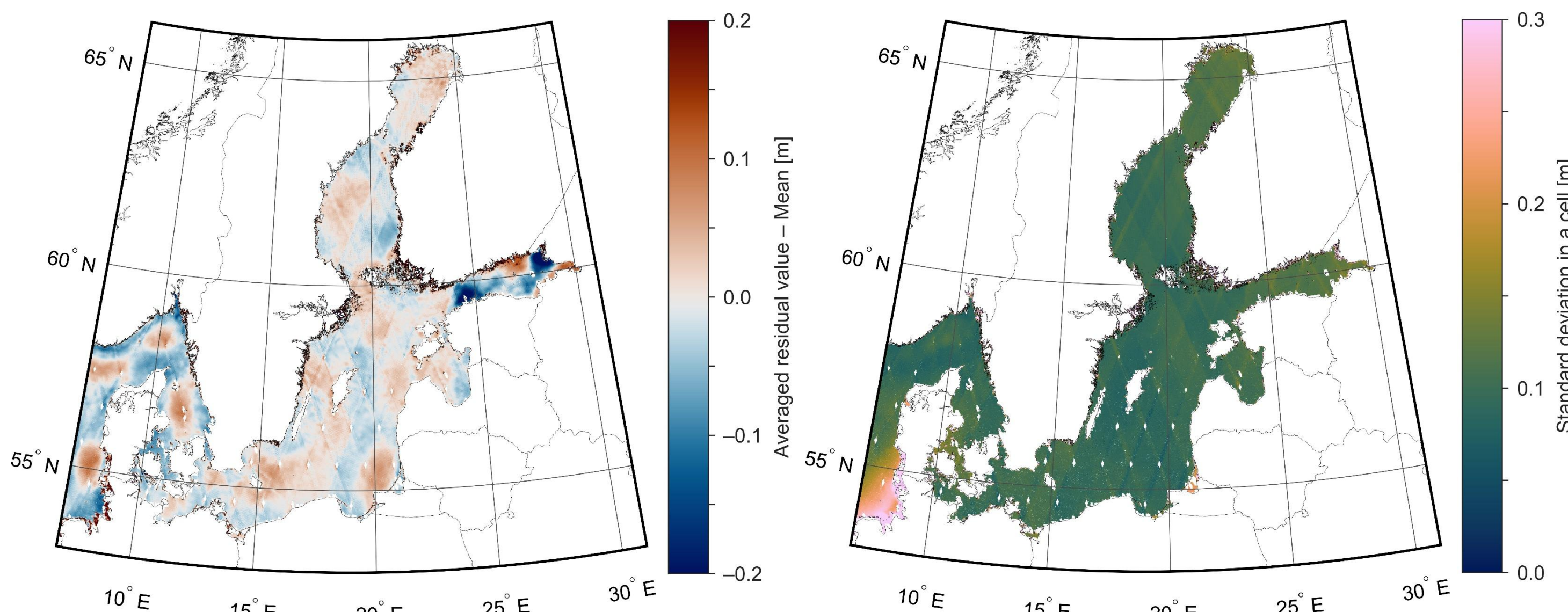


Figure 1. Averaged residual values r_{ij}^{avg} (left), after subtraction of the mean residual, and cell-based standard deviation estimates σ_{ij}^{avg} (right). Both quantities are shown after removal of outliers.

Acknowledgements

This research was funded by the Estonian Research Council grant no. PUTJD1247.

Relevant references

- Copernicus Marine Service. (2026). Global Ocean Physics Analysis and Forecast. Copernicus Marine Service. DOI: 10.48670/moi-00016.
- Förste, C.; Bruinsma, S. L.; Abryksov, O.; Lemoine, J.-M.; Marty, J. C.; Flechtner, F.; Balmino, G.; Barthelmes, F.; Biancale, R. (2014). EIGEN-6C4 the latest combined global gravity field model including GOCE data up to degree and order 2190 of GFZ Potsdam and GRGS Toulouse. GFZ Data Service. DOI: 10.5880/igem.2015.1.
- Jet Propulsion Laboratory. (2025). SWOT product description: L2_LR_SSH (document no. JPL D-56407, revision C). NASA Jet Propulsion Laboratory.
- Kearsley, W. (1977). Non-stationary estimation in gravity prediction problems. Ohio State University, Department of Geodetic Science, report #256, 63 pp.
- Moritz, H. (1980). Advanced physical geodesy. Wichmann, Karlsruhe.
- Schwabe, J.; Varbla, S.; Ågren, J.; Teitsson, H.; Elmann, A.; Liebsch, G.; Forsberg, R.; Strykowski, G.; Bilker-Koivula, M.; Liepiņš, I.; Paršeliūnas, E.; Keller, K.; Omang, O. C. D.; Vestøl, O.; Kaminski, J.; Wilde-Piörko, M.; Szelachowska, M.; Pyrcia, K.; Somla, J.; Westfeld, P.; Hammakint, T.; Olsson, P.-A.; Förste, C.; Ince, E. S. (2026). The development of the unified Baltic Sea Chart Datum 2000 (BSCD2000) height transformation grid: an unprecedented example of height system unification based on state-of-the-art marine geoid modelling. Journal of Geodesy, 100, #65. DOI: 10.1007/s00190-026-02096-z.
- Shaw, L.; Paul, I.; Henrikson, P. (1969). Statistical models for the vertical deflection from gravity-anomaly models. Journal of Geophysical Research, 74, 4259-4265. DOI: 10.1029/JB074i017p04259.
- Stiles, B.; Dubois, P. (2023). SWOT algorithm theoretical basis document: L2_LR_SSH (document no. JPL D-105502). NASA Jet Propulsion Laboratory.
- SWOT. (2024). SWOT Level 2 KaRIn low rate sea surface height data product, version C. NASA Physical Oceanography Distributed Active Archive Center. DOI: 10.5067/SWOT-SSH-2.0.

METHOD (GRIDDED SURFACE)

- The empirical autocovariance curve is derived from averaged residual values r_{ij}^{avg} , denoted for convenience by a single index α (or β for a second estimate in a pair) corresponding to unique coordinates, such that $r_\alpha = r_{ij}^{avg} - \mu[r_{ij}^{avg}]$, where $\mu[\cdot]$ denotes the mean over all estimates r_{ij}^{avg} , ensuring that the transformed data are centred. For each pair of averaged residual values (α, β) , $\alpha \neq \beta$, a pairwise covariance estimator is calculated as $C_{\alpha\beta} = r_\alpha r_\beta$. The modulated empirical curve (see the red scatter below) is then computed as:

$$C_{emp}(l_b) = \frac{1}{n_b} \sum_{(\alpha, \beta) \in B_b} w_{\alpha\beta} C_{\alpha\beta}$$

where n_b is the number of contributing unique pairs in the discrete distance interval b (forming a set B_b), and covariance distance l_b is defined as the midpoint of the corresponding distance interval (i.e., the bin centre measured from zero separation).

- The modulation weights are defined as $(\sigma_\alpha = \sigma_{ij}^{avg})$:

$$w_{\alpha\beta} = \left[\left(\frac{\sigma_\alpha}{r_\alpha} \right)^2 + \left(\frac{\sigma_\beta}{r_\beta} \right)^2 \right]^{-1/2}$$

The quality-dependent modulation suppresses the influence of low signal-to-noise ratio pairs, thereby enabling robust signal extraction from the initial noisy data. The approach also avoids over-smoothing of the empirical autocovariance curve compared to standard inverse-variance weighted averaging.

- Correlation length $X_{1/2}$ is derived through least-squares fitting of a theoretical model with unconstrained signal variance (see the red line below).

- A weighted signal variance C_0 is computed using r_{ij}^{avg} with weights defined as $1/(\sigma_{ij}^{avg})^2$ to restore the correct covariance scale (see the magenta line below).

- The obtained empirical autocovariance curve is best represented by the exponential covariance model $C_{model}(l_b) = C_0 e^{-l_b/\gamma}$, where $\gamma = (-\ln 0.5)^{-1} X_{1/2} \approx 1.443 X_{1/2}$.

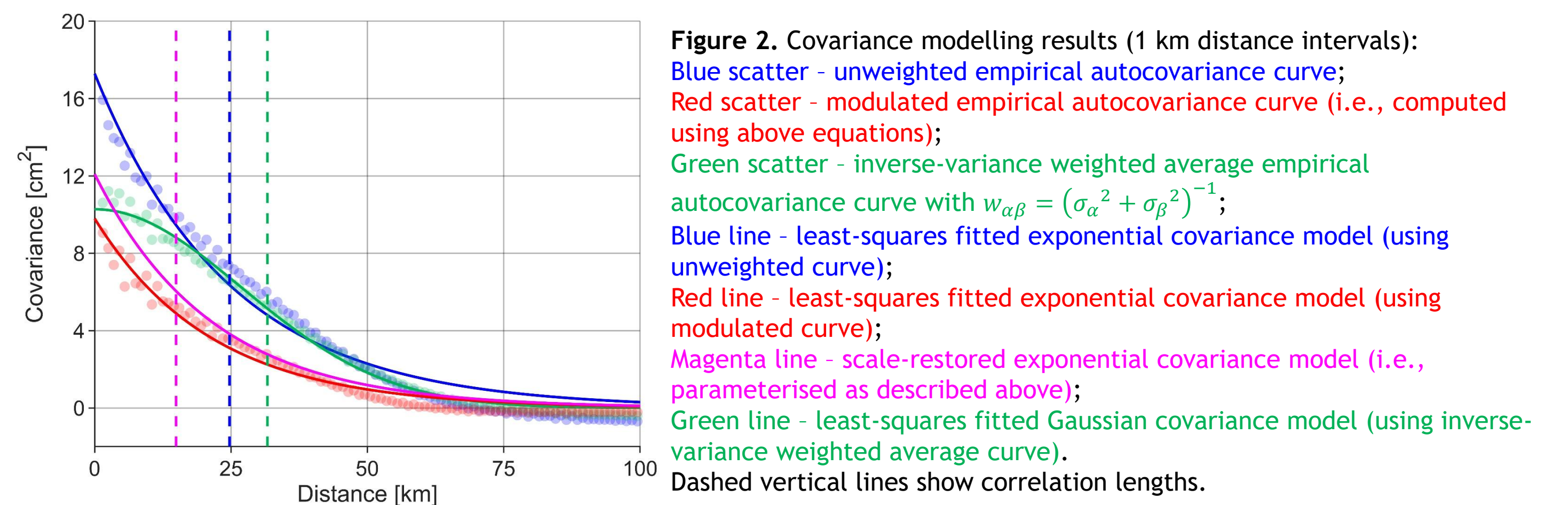


Figure 2. Covariance modelling results (1 km distance intervals): Blue scatter - unweighted empirical autocovariance curve; Red scatter - modulated empirical autocovariance curve (i.e., computed using above equations); Green scatter - inverse-variance weighted average empirical autocovariance curve with $w_{\alpha\beta} = (\sigma_\alpha^2 + \sigma_\beta^2)^{-1}$; Blue line - least-squares fitted exponential covariance model (using unweighted curve); Red line - least-squares fitted exponential covariance model (using modulated curve); Magenta line - scale-restored exponential covariance model (i.e., parameterised as described above); Green line - least-squares fitted Gaussian covariance model (using inverse-variance weighted average curve). Dashed vertical lines show correlation lengths.

- Averaged residual values r_α (in vector $\mathbf{r}$) form the basis for LSC-based gridding:

$$\mathbf{r}^{LSC} = \mathbf{C}_{st}(\mathbf{C}_{tt} + \mathbf{C}_{nn})^{-1} \mathbf{r}$$

where $\mathbf{C}_{st}$ is the cross-covariance matrix of the predicted (r^{LSC}) and known (r_a) signals, and $\mathbf{C}_{tt}$ and $\mathbf{C}_{nn}$ are the covariance matrices of known signal and observation errors (σ_{ij}^{avg}), respectively. The matrices $\mathbf{C}_{st}$ and $\mathbf{C}_{tt}$ are derived using the scale-restored exponential covariance model (see the magenta line above).

- The geoid-consistent height reference surface is determined by adding back the mean of all averaged residual values r_{ij}^{avg} and restoring the initially removed reference field:

$$N^{LSC}(\varphi, \lambda) = r^{LSC}(\varphi, \lambda) + \mu[r_{ij}^{avg}] + N_{REF}(\varphi, \lambda)$$

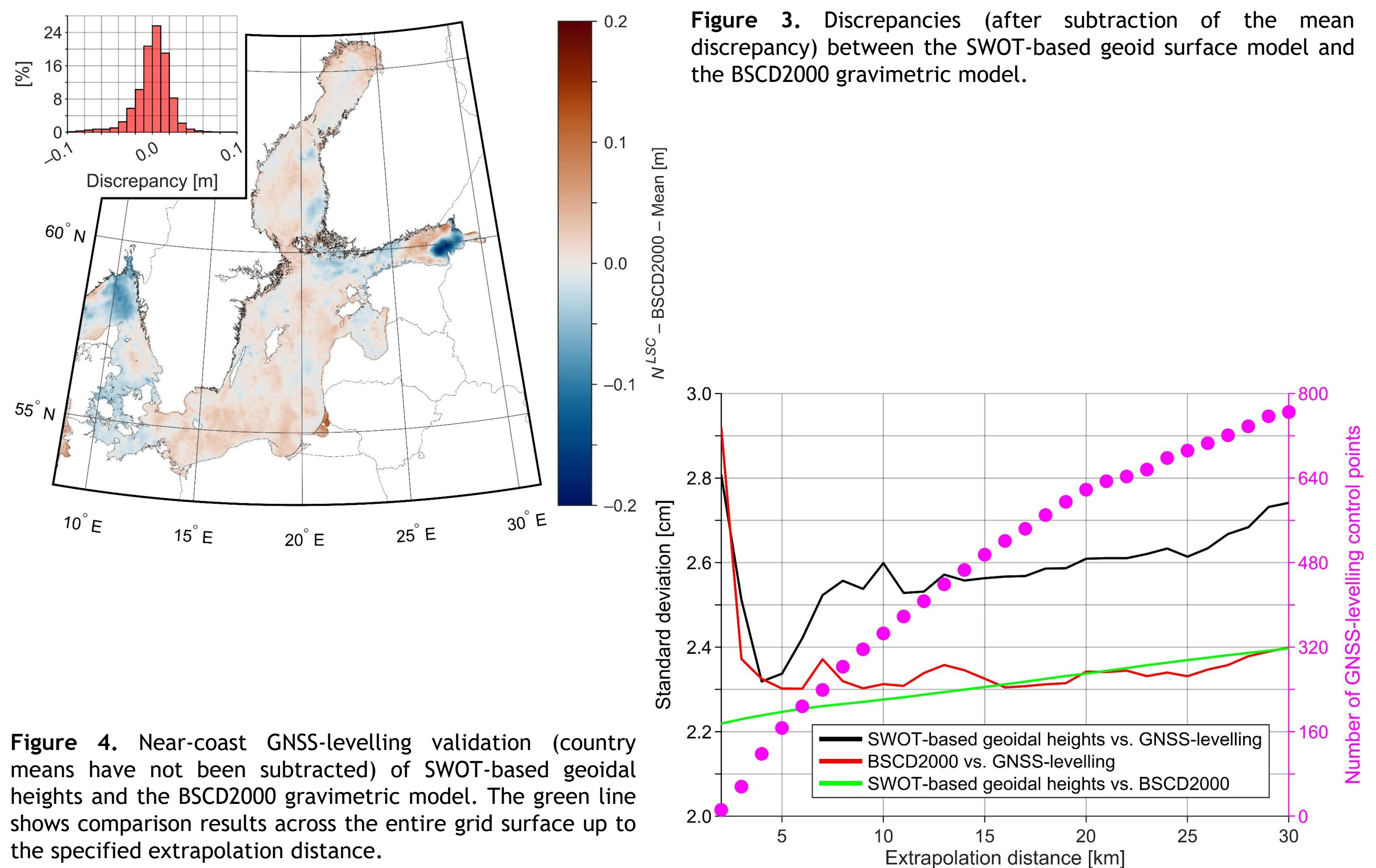


Figure 3. Discrepancies (after subtraction of the mean discrepancy) between the SWOT-based geoid surface model and the BSCD2000 gravimetric model.

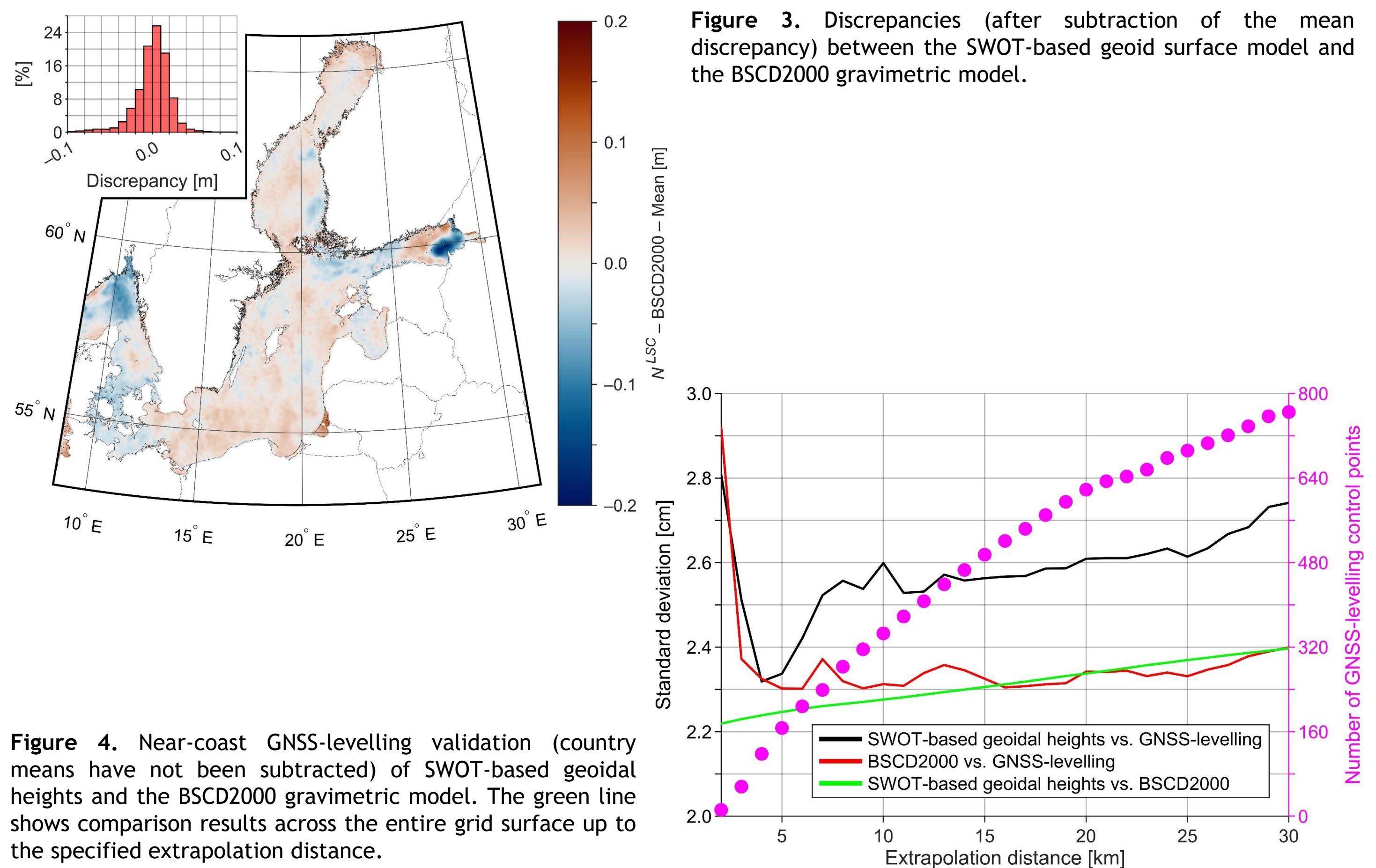


Figure 4. Near-coast GNSS-levelling validation (country means have not been subtracted) of SWOT-based geoidal heights and the BSCD2000 gravimetric model. The green line shows comparison results across the entire grid surface up to the specified extrapolation distance.

CONCLUSIONS

- The resulting SWOT-based geoid surface model agrees well with the gravimetric BSCD2000 quasigeoid model, with most differences within ± 3 cm and a standard deviation of 2.2 cm.
- Validation using coastal GNSS-levelling control points yields a standard deviation of 2.3 cm within an extrapolation distance of 5 km, indicating good consistency with the BSCD2000 model in the coastal zone.
- The results demonstrate that SWOT altimetry observations provide a viable geometric alternative to conventional gravimetric marine geoid determination. The determined geometric model can also support offshore data quality control and geoid modelling diagnostics (e.g., in the problematic eastern Gulf of Finland).