

Reply to RC1 for the paper titled: Sea level reconstruction reveals improved separations of regional climate and trend patterns over the last seven decades.

Author(s): Shengdao Wang et al.

MS No.: essd-2025-251

MS type: Data description paper

RC1 general comments:

The manuscript presents a global sea level reconstruction spanning 1950–2022 at a $1^\circ \times 1^\circ$ resolution, using a combined dataset of tide gauge records (corrected for vertical land motion) and satellite altimetry, implemented through an enhanced CSEOF-OI framework incorporating EOF decomposition. The authors emphasize improved separation of climate modes (e.g., ENSO, PDO) and long-term trends.

While the study contributes to an important topic and provides a potentially valuable dataset to the sea level and climate research community, several key aspects of the methodology and validation require further clarification. In particular, the validation appears limited to coarser $5^\circ \times 5^\circ$ spatial scales, and insufficient attention is given to the accuracy of reconstructed data in the pre-altimetry era (especially before 1993) and in open-ocean regions lacking tide gauge constraints. These issues, along with others detailed below, should be addressed to ensure the robustness and usability of the dataset.

Dear Reviewer and Editor,

Thank you for your thoughtful and detailed comments. We have carefully considered and answered each point and revised the manuscript accordingly. In response to your main concerns, (1) we added a schematic for modified CSEOF–OI reconstruction workflow into the supplementary information; (2) expanded the validation from $5^\circ \times 5^\circ$ to the native $1^\circ \times 1^\circ$ grid and revised comparison in the Figure 6e–g accordingly and we also further provide gridded $1^\circ \times 1^\circ$ uncertainty (error) fields for each time step in the updated data product link and (3) provide two additional figures and one table to help better answer the questions (RC1_Fig1, RC1_Fig2, RC1_Table1). Below, we respond to your questions and comments point by point. For readability, our responses are in **blue**, and the corresponding manuscript changes are shown in **red**. We believe that these revisions have substantially strengthened the paper based on your valuable suggestions.

On behalf of all authors

Best regards,

Shengdao Wang

RC1-Q1: The manuscript reports strong correlations ($r = 0.87$ and $r = 0.75$) between reconstructed principal components and ENSO/PDO indices. However, the interpretation as “excellent agreement” should be moderated, and the authors should clarify whether statistical significance tests were performed and how preprocessing (e.g., detrending) was handled.

Reply: Thanks for the suggestion. We moderated the word from “excellent agreement” to “good agreement”. We applied statistical significance tests of the PC-index correlations, following Ebisuzaki (1997), using a phase-randomized surrogate test to account for potential autocorrelation in the datasets. And we added citations to the related literature to the manuscript. Because neither the PCs nor the climate indices exhibited apparent residual trend, we did not apply any additional detrending. With $N = 10000$ surrogates, we obtain correlation between annual smoothed PC3 and respective ENSO (Niño-3.4) index $r = 0.87$ and $p = 0.0001$; PC4 and PDO show $r = 0.75$, $p = 0.0001$. The p -values come from the Ebisuzaki phase-randomization test (null: zero correlation between the two series). These results confirm that the correlations are statistically significant. We added this clarification to the revised manuscript.

Detailed changes in the manuscript:

- In line 511, we added: The Climate index scale matched with PCs, **no additional detrending was applied.**
- In line 532, we added: **Correlation significance was evaluated with a phase-randomized surrogate test (Ebisuzaki, 1997; $N = 10000$). The annually smoothed PC3–Niño-3.4 correlation is statistically significant ($p = 0.0001$).**
- And in lines 542 to 543 added: The associated smoothed PC4 (Figure 7l) exhibits a robust correlation ($r = 0.75$, $p = 0.0001$ **at the surrogate tests**) with the annually smoothed PDO index
- Newly added Ebisuzaki (1997) reference at Line 841: **Ebisuzaki, W.: A Method to Estimate the Statistical Significance of a Correlation When the Data Are Serially Correlated, J. Climate, 10, 2147–2153, [https://doi.org/10.1175/1520-0442\(1997\)010<2147:AMTETS>2.0.CO;2](https://doi.org/10.1175/1520-0442(1997)010<2147:AMTETS>2.0.CO;2), 1997.**

RC1-Q2: In Fig. 1(a), how many tide gauge stations are included in total? Since the data record lengths vary across stations, the authors should also clarify the proportion of stations within each time span category as shown in the figure. In Fig. 1(b), the authors present the percentage distribution of tide gauge records covering the 1950–2022 period. However, it would be helpful to specify the actual number of stations corresponding to each percentage category.

Reply: We appreciate the reviewer’s careful reading. There is a total of 1537 tide gauge stations shown in Fig.1 (a and b), and we additionally added the number & proportion information of each type of tide gauge station in their Figure captions.

Detailed changes in the manuscript:

- Figure 1 caption in Lines: 114 to 116 revised into: Figure 1 (Caption revised, figure no changes). Global distribution of PSMSL (Permanent Service for Mean Sea Level) tide-gauge stations (a total of 1537 stations). (a) Stations colored by record duration (in full years). Ratio of stations in each duration range: ≤ 30 years (787/1537), 30–45 years (258/1537), 45–60 years (221/1537), 60–75 years (118/1537), 75–90 years (51/1537), 90–105 years (41/1537), 105–120 years (26/1537), ≥ 120 years (35/1537). (b) Percentage of each station’s record available within 1950–2022. Ratio of stations in each completeness percentage range: $\leq 40\%$ (801/1537), 40–50% (156/1537), 50–60% (114/1537), 60–70% (117/1537), 70–80% (117/1537), 80–90% (78/1537), $\geq 90\%$ (154/1537). Base map image: NASA (National Aeronautics and Space Administration) Blue Marble.

RC1-Q3: The manuscript refers to the data time span as both “January 1950 to December 2021” and “January 1950 to January 2022” in different sections. The authors should ensure consistency and clarify the exact temporal coverage of the reconstructed dataset.

Reply: Thank you for pointing this out. The apparent confusion arises from Line 134, where the gap-filling simulation was run with an a priori selected data span January 1950–December 2021 (864 months), using PSMSL tide-gauge data downloaded in October 2022. At that time, most of the tide gauge records were still not updated to January 2022, so to maximize near-complete gauge record numbers, we did not include January 2022 in the gap-filling simulation. The simulation was designed to test the impact of different types of gaps in the long-term monthly tide gauge records. So, adding January 2022 only appends one more month to the 72-year records, and the comparative ranking of methods remains unchanged, i.e., the Regularized EM method on tide gauge gap fillings outperforms PPCA and AR modeling. To keep the test setup constant and reproducible, we used 48 PSMSL gauges in the initial download for that simulation.

For the final $1^\circ \times 1^\circ$ monthly reconstructed sea level product—because it is the data product of this study—to maximize temporal coverage. Therefore, before manuscript submission, we finalized our reconstructed sea level by applying the 2024 PSMSL download. In that release, most gauges used in the reconstruction do include January 2022; consequently, we reconstruct the gridded sea-level product spans January 1950–January 2022.

Detailed changes in the manuscript:

- To clarify and avoid potential confusion, we changed Line 134 to 135 into: The simulation study selected 48 near-complete (average data-missing rate 1.1%) global long-term tide-gauge records (864 months; span from beginning of 1950 to end of 2021) from the PSMSL Revised Local Reference (RLR) monthly product to evaluate the performance of different gap-filling methods for dealing with gaps in multi-decadal gauge data.

RC1-Q4: Lines 141–142 refer to Figs. S1–S3, and line 165 mentions Fig. S4. However, these figures are not provided in the manuscript. The authors should ensure that all referenced supplementary figures are included at the end of the paper.

Reply: Thank you for flagging this. The supplementary figures S1–S4 were provided at the initial submission as part of the Supplementary Information available via the journal’s link (<https://essd.copernicus.org/preprints/essd-2025-251/essd-2025-251-supplement.pdf>). To avoid any confusion, we will further attach this supplementary information to the revised manuscript.

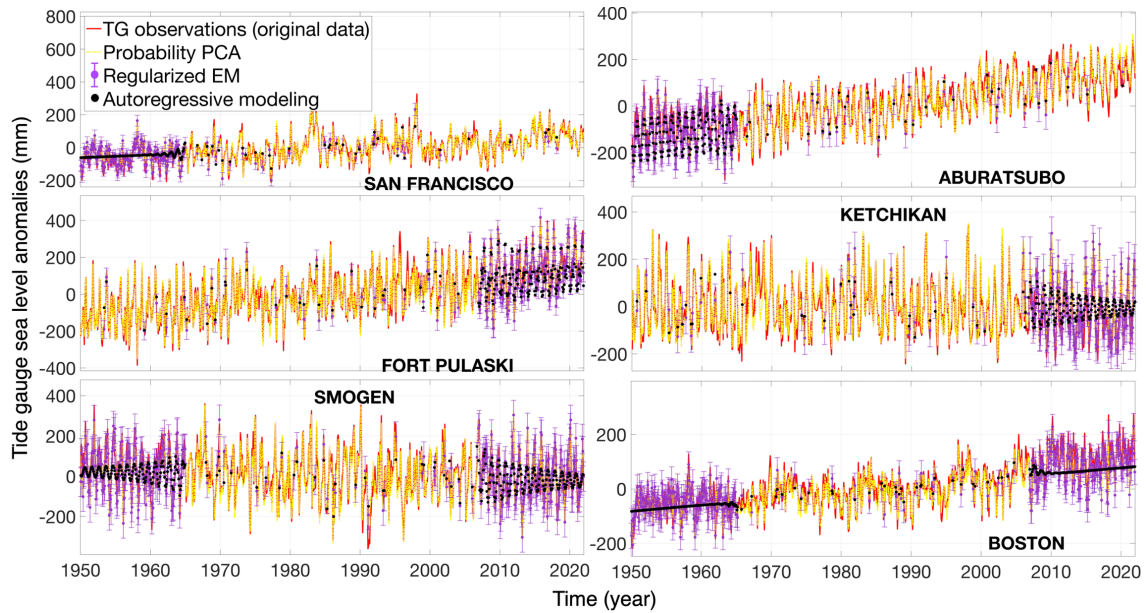
RC1-Q5: The manuscript states that missing data in tide gauge records were filled using AR, PPCA, and EM methods, with EM identified as the optimal approach based on validation experiments. However, it is unclear how the authors handled records with substantial data gaps at the beginning or end of the time series. What was the quality of the gap-filling in such cases? As Fig. 1(b) indicates, some stations have more than 50% missing data — were these records also gap-filled and subsequently used in the sea level reconstruction? Further clarification is needed regarding the treatment and reliability of heavily gapped records.

Reply: Regularized EM (Schneider, 2001) estimates missing values using the full network’s space–time covariance with data-adaptive weights. In our tide-gauge gap filling, estimates are therefore not driven by any single station; the method leverages all available stations and temporal dependencies. Prior work demonstrates skill for long start-/end-of-record and discontinuous gaps—for example, filling large start-of-record gaps and validating against the overlapping instrumental period (Mann et al., 2007), and Antarctic temperature series with large edge gaps (Steig et al., 2009).

Beyond the tide gauge gaps simulations illustrated in Figure 2, Table 1, and Figs. S1–S3 in Supplementary Information, we additionally designed a stress test for the ‘extreme’ cases, which reviewers mentioned. From 48 near-complete 72-year (864-month) tide-gauge series in Figure 2, we randomly selected six gauges with created three extreme edge-gap cases: (i) two gauges with a 15-year (180-month) start gap (San Francisco, Aburatsubo); (ii) two with a 15-year end gap (Fort Pulaski, Ketchikan); and (iii) two with 15-year gaps at both ends (Smögen, Boston). On top of each edge gap, we inserted an additional 5 years (60 months) of random internal missing data.

We compared RegEM (Regularized EM), PPCA, and AR using correlation (r) and RMSE between their predicted results and respective real observations (results detailed in RC1_Fig1 and RC1_Table1). PPCA reconstructs the entire series, so we show its full output; RegEM and AR can only predict the missing intervals while leaving observed values unchanged (illustrated in RC1_Fig1).

RC1_Table1 shows that even under these extreme cases, RegEM can still maintain agreement between simulated gap-filled gauge records with the truth observations under extreme gaps at $r \approx 0.97$ – 0.99 across six test gauges and achieves the lowest RMSE among the three methods.



RC1_Fig1. Comparisons of three tide-gauge gap-filling methods—Autoregressive (AR) modeling, Probabilistic PCA (PPCA), and Regularized EM. We test “extreme” edge-gap cases: beginning gaps of 180 months (top two panels), end gaps of 180 months (middle two), and both ends with 180-month gaps (bottom two). Each record also contains an additional 60-month random internal gap outside the large edge gap. The red line shows the “true” observations. PPCA reconstructs the entire time series (full visualization), whereas AR and EM display only the filled gap/missing segments, with Regularized EM showing 95% confidence intervals of each predicted missing value.

RC1_Table1. Comparing the performance of the three data gap-filling methodologies based on the average Pearson correlation coefficient (r) and Root Mean Square Errors (RMSE) between filled missing records and true observations in the extreme gaps simulation experiments.

Stations	San Francisco	Aburatsubo	Fort Pulaski	Ketchikan	Smogen	Boston
Parameters						
Correction coefficient with TG observations (r)	$r = 0.98$ (RegEM) $r = 0.96$ (PPCA) $r = 0.93$ (AR)	$r = 0.99$ (RegEM) $r = 0.97$ (PPCA) $r = 0.91$ (AR)	$r = 0.99$ (RegEM) $r = 0.98$ (PPCA) $r = 0.93$ (AR)	$r = 0.99$ (RegEM) $r = 0.97$ (PPCA) $r = 0.92$ (AR)	$r = 0.98$ (RegEM) $r = 0.97$ (PPCA) $r = 0.83$ (AR)	$r = 0.97$ (RegEM) $r = 0.94$ (PPCA) $r = 0.71$ (AR)
Root-mean-squared error (RMSE, mm)	18.60 mm (RegEM) 24.78 mm (PPCA) 31.25 mm (AR)	10.52 mm (RegEM) 18.31 mm (PPCA) 29.65 mm (AR)	14.73 mm (RegEM) 21.47 mm (PPCA) 41.21 mm (AR)	15.92 mm (RegEM) 25.87 mm (PPCA) 45.31 mm (AR)	24.93 mm (RegEM) 29.53 mm (PPCA) 64.67 mm (AR)	18.31 mm (RegEM) 22.72 mm (PPCA) 43.87 mm (AR)

For robustness, we additionally repeated the experiment 500 times: in each run, six gauges were randomly sampled from the set of 48 (two with a start gap, two with an end gap, and two with gaps at both ends), and each gauge also received an additional random 5-year missing. Over 500 runs, RegEM achieved an average correlation of 0.97 ± 0.012 between the gap-filled series and the true gauge series (PPCA: 0.95 ± 0.016 ; AR: 0.88 ± 0.015) and the lowest RMSE (23.51 ± 3.52 mm; PPCA: 27.97 ± 3.00 mm; AR: 52.73 ± 7.71 mm), demonstrating superior performance under extreme gap scenarios.

Across the 225 gauges used in the reconstruction, the average missing ratio during Jan 1950–Jan 2022 is 12%, and we apply a missing-ratio threshold $\leq 40\%$. Two gauges—Sheerness (UK) and Whangārei Harbour (New Zealand)—were retained as justified exceptions because nearby

stations within 120 km provide $\geq 80\%$ overlap measurements for each ≥ 5 -year gap, which allows stable RegEM fill missing as the method emphasizes using space–time covariance and data-adaptive weights to estimate missing values.

More importantly, we already account for errors from gap filling in the final reconstruction. Specifically, we run a Monte Carlo ensemble that propagates gap-filling-induced errors through the full pipeline and use the ensemble spread at each time step as the source’s contribution to reconstructed sea level uncertainty, as detailed in section 2.3.

RC1-Q6: The sentence “GIA models have been extensively used to harmonize measurements between altimetry and tide gauge sea level” is misleading. GIA corrections are primarily used to account for long-term vertical land motion, not to harmonize the two types of measurements. The authors should revise this statement to more accurately reflect the role of GIA models.

Reply: We appreciate the reviewer’s clarification, and we agree. GIA corrections are primarily applied to account for the GIA-related long-term vertical land motion at tide-gauge sites. The confusion arose from our use of the word “harmonize.” To avoid this ambiguity, we have rewritten the sentence in the revised manuscript.

Detailed changes in the manuscript:

- We revised the sentence in Line 198 into: In many prior studies on sea level reconstructions, GIA models are used to correct the GIA-related long-term vertical land motion at tide-gauge sites, thereby making tide-gauge relative sea level more consistent with geocentric sea level from satellite altimetry.

RC1-Q7: The manuscript discusses the limitations of GIA models and emphasizes that GIA is not the sole contributor to vertical land motion at tide gauge locations. However, it remains unclear how GIA corrections were specifically implemented in the reconstruction. The authors should explicitly state which GIA model was used, how the corrections were applied, and whether additional vertical land motion sources (e.g., tectonics, anthropogenic subsidence) were considered or corrected for in the analysis.

Reply: Thank you for the comment. GIA (Glacier Isostatic Adjustment) is discussed only as background regarding VLM (vertical land motion) correction/estimation at tide gauge locations. And for the gauges which were used for our sea level reconstruction product, we did not apply the GIA model to account for VLM; instead, the method we applied is deriving rates from the satellite-altimetry minus tide-gauge (SA–TG) as detailed in Line 241 to Line 246.

Detailed changes in the manuscript:

- For clarity and avoid confusing, we will add one sentence before Line 198: “Contemporary approaches for estimating vertical land motion (VLM) at tide-gauge sites

include applying glacial isostatic adjustment (GIA) models, using Global Navigation Satellite System (GNSS) observations, and deriving rates from the satellite altimetry minus tide gauge (SA-TG) difference; we will firstly briefly discuss these methods and subsequently select the approach that is appropriate for our study.”

RC1-Q8: The discussion on vertical land motion (VLM) from lines 209 to 226 reads more like a general literature review and would be more appropriately placed in the Introduction section. The Methods section should focus on describing the specific data sources and procedures used in this study.

Reply: We appreciate the reviewer’s suggestion. Our intent with the original text was to provide additional detail to help readers know tide-gauge vertical land motion (VLM) correction. To better focus, we will therefore shorten the background and initially mention it in the Introduction, while keeping Section 2.2 focused on the specific method we actually used (SA-TG for estimating VLM at tide gauges). VLM correction at tide gauges is one of the necessary technical components of sea-level reconstruction; we suggest keeping a concise, methods-oriented description in Section 2.2. In the Introduction, we will add one bridging sentence near lines 69–71 to orient readers to the three common VLM approaches (GIA models, GNSS observations, and SA–TG) and point them to Section 2.2 for details.

Detailed changes in the manuscript:

- New sentence for the Introduction, add in red (near lines 69–71)
The contemporary methodology integrates temporally extensive but spatially limited records from tide gauges—whose vertical land motion is accounted via glacial isostatic adjustment (GIA) models, Global Navigation Satellite System (GNSS) observations, or the satellite altimetry minus tide-gauge difference (see Section 2.2)—and satellite altimetry measurements that provide extensive spatial coverage and resolution, but with limited temporal duration.
- Revision for Lines 209–226 (concise for better focus)

Global Navigation Satellite System (GNSS) can provide VLM estimates at collocated gauges, but in practice, its utility is limited by the scarcity of long, stable records (Wöppelmann et al., 2019) and influenced by long-period geophysical signals (Santamaría-Gómez & Mémin, 2015; Santamaría-Gómez et al., 2017). Additionally, in some regions, even minor geographic misalignments (a few kilometers) between the GNSS receiver and tide gauge can cause GNSS-estimated vertical land motion to inadequately represent gauge location motions due to the high spatial variability (Bevis et al., 2002; Oelmann et al., 2021). Precise leveling provides an accurate tie, but it is labor-intensive and rarely maintained.

Other methods to estimate VLM have used the satellite altimetry minus tide gauge (SA–TG) approach (Nerem and Mitchum, 2002; Kuo et al., 2004, 2008; Ray et al., 2010; Wan,

2015; Wöppelmann and Marcos, 2016; Oelsmann et al., 2021). In particular, Kuo et al. (2004) and Kuo et al. (2008) applied a network adjustment to reduce VLM error estimates to $<0.5 \text{ mm yr}^{-1}$ in semi-enclosed seas and lakes. Wan (2015) jointly solved for geocentric sea-level trend and VLM at global tide gauges. Wöppelmann and Marcos (2016) further showed that accurate VLM estimation via SA–TG requires a common data span of ≥ 20 years between altimetry and tide gauge data—a criterion that has been met.

RC1-Q9: The sentence in lines 227–228 (“An issue that might be regarded as a mere matter of language is examined here; however, imprecise terminology can give rise to subtle yet conceptual misunderstandings”) is vague in both meaning and context. It is unclear what specific issue the authors are referring to in this section. The authors should clarify the intended point or consider removing or relocating the sentence for better coherence.

Reply: Thank you for pointing this out. The purpose of the sentence is to help emphasize the upcoming several sentences; removing it does not affect the logic, results, or conclusions. Therefore, we have removed it.

Detailed changes in the manuscript:

- We agree it may disrupt the flow, so we will remove it. Lines 227–228: ~~An issue that might be regarded as a mere matter of language is examined here; however, imprecise terminology can give rise to subtle yet conceptual misunderstandings.~~

RC1-Q10: What is the spatial resolution of the AVISO Level 4 monthly gridded sea level product used in this study? How are the sea level changes at the tide gauge locations derived from the gridded data? Specifically, what spatial matching or interpolation methods are applied to relate the gridded data to the tide gauge positions?

Reply: For the AVISO product and resolution. We use the AVISO Level-4 monthly gridded sea-level product (Jan 1993–Dec 2021) at $\frac{1}{4}^{\circ} \times \frac{1}{4}^{\circ}$ ($\sim 0.25^{\circ}$) resolution and resample to $1^{\circ} \times 1^{\circ}$ to target regional-to-global scales and consistent with prior studies on similar topics (Church & White, 2011; Calafat et al., 2014; Dangendorf et al., 2019), then perform the $1^{\circ} \times 1^{\circ}$ sea-level reconstruction by combining with tide gauges. The reconstructed sea level is the 72 years of monthly sea level at each $1^{\circ} \times 1^{\circ}$ grid cell. Following these three studies (Church & White, 2011; Calafat et al., 2014; Dangendorf et al., 2019), we do not spatially interpolate the gridded sea level to the tide-gauge locations. Instead, before sea level reconstruction, each gauge is assigned to the nearest $1^{\circ} \times 1^{\circ}$ grids, and gauges used in the reconstruction process are those located within 100 km of the nearest grid point (as described in Lines 288–289). And we do not further differentiate any sea-level differences caused by the distance between the tide-gauge locations and their nearest 1° grid cells.

Note: Q11 and Q17 are related. Because the response to Q17 partly addresses Q11, we treat them together—answering Q17 first, then Q11.

RC1-Q17: Line 265 mentions that GNSS-derived and SA–TG-derived VLM time series are not consistent. However, the manuscript does not clarify the temporal coverage of GNSS observations used for validation. What is the time period of the GNSS-derived VLM estimates across the 253 tide gauge sites? Are these periods consistent across stations? Moreover, temporal inconsistencies between GNSS and SA–TG estimates may lead to biased trend comparisons, especially if GNSS observations cover shorter or non-overlapping periods with significant non-linear land motion. This potential impact should be addressed and quantified. We note that comparisons of vertical land motion should be conducted over consistent time periods; otherwise, the comparison becomes fundamentally invalid and cannot reliably assess the agreement between the two methods.

Reply: Thank you for the questions. For clarity, we first note a commonly recognized background point: converting long-term tide-gauge (e.g., 50–100 year) relative sea level (RSL) to absolute/geocentric sea level (ASL) requires accounting for vertical land motion (VLM). Three approaches are commonly used—GIA models, collocated/nearby GNSS, and the satellite altimetry minus tide gauge (SA–TG) estimated VLM rate—each with known assumptions and limitations (Peltier et al., 2015; Santamaría-Gómez et al., 2012; Wöppelmann and Marcos, 2016; He et al., 2025). More importantly, most tide-gauge sites do not have VLM observations spanning the full 50–100 years; therefore, the conversion (long-term RSL to long-term ASL) relies on VLM solutions that are not measured across the entire TG period (Church and White, 2011; Santamaría-Gómez et al., 2012; Wöppelmann and Marcos, 2016; Dangendorf et al., 2017; Bruni et al., 2022). To limit errors from short records, the community applies different minimum record-length criteria for GNSS and SA–TG VLM because these two methods hold different noise spectra and confounders (Williams et al., 2004; Wöppelmann and Marcos, 2016). For GNSS-derived VLM at tide gauges, some GPS solutions adopt a minimum record length of about 3.5 years (e.g., Wöppelmann et al., 2009; Hammond et al., 2021), whereas SA–TG typically requires ≥ 20 years of altimetry and TG overlap for robust long-term rates (e.g., Wöppelmann and Marcos, 2016).

With this background, we address **Q17** in detail: The GNSS-based VLM rates used for validation are from Hammond et al. (2021). Their solution applies standard screening (record length ≥ 3.5 years) and models annual/semiannual terms and offsets/jumps to mitigate non-linear behavior (e.g., steps). Record lengths at tide gauge sites are heterogeneous (~ 3.5 – 26.5 years). The data duration of the GNSS stations shown in Figure 3 spans from October 1994 to December 2020, with a median record length of 9.93 years (mean 10.42 years). By contrast, the SA–TG method requires a long altimetry–tide-gauge overlap; community practice recommends $\geq \sim 20$ years for robust long-term rates. We therefore adopt a uniform 29-year window (1993–2021) for SA–TG. We do not recompute SA–TG over each site’s GNSS span because the two approaches have different minimum-length requirements. Previously, a lot of studies compared GNSS and SA–TG based VLM did not force identical windows (e.g., Wöppelmann & Marcos, 2016; Oelsmann et al., 2021); imposing short GNSS length windows on SA–TG would violate recommended practice and yield unstable SA–TG estimates, as most GNSS don’t have data length over 20 years. The key is that each estimate needs to meet its own validity data span criterion.

Detailed changes in the manuscript:

- We added a sentence at line 256 to clarify this: Note that GNSS and SA–TG based VLM rates have different minimum effective record length requirements—GNSS typically ≥ 3.5 years (e.g., Wöppelmann et al., 2009; Hammond et al., 2021) versus SA-TG $\geq \sim 20$ years (Wöppelmann & Marcos, 2016); all solutions shown in Figure 3 satisfy these criteria.
- Newly added Wöppelmann et al., 2009 reference: Wöppelmann, G., Letetrel, C., Santamaria, A., Bouin, M. -N., Collilieux, X., Altamimi, Z., Williams, S. D. P., and Miguez, B. M.: Rates of sea-level change over the past century in a geocentric reference frame, *Geophysical Research Letters*, 36, 2009GL038720, <https://doi.org/10.1029/2009GL038720>, 2009.

Next, we address the details in Q11

RC1-Q11: The comparison between SA–TG-derived and GNSS-derived VLM estimates shows a median difference of 0.88 mm/yr ($r = 0.86$) across 253 sites. Even in the subset with the smallest GNSS uncertainties, the median difference remains 0.64 mm/yr. These differences are non-negligible and may influence long-term sea level trend estimates. The authors should discuss the implications of such discrepancies on their reconstruction results. Additionally, the manuscript does not report the average (mean) VLM rates estimated separately by SA-TG and GNSS methods over the 253 sites; this information should be provided to better understand the characteristics of both estimates.

Reply: We agree with the reviewer that the smallest-uncertainty subset still shows a median difference of 0.64 mm year⁻¹. The purpose of this value is (i) evidence of convergence—the median absolute difference decreases monotonically from 0.88 $\rightarrow$ 0.81 $\rightarrow$ 0.64 mm yr⁻¹ while the correlation increases from $r = 0.86 \rightarrow 0.89 \rightarrow 0.95$ as GNSS quality improves—and as an external comparison only. In our sea level reconstruction, we do not replace SA-TG estimated VLM rates with GNSS VLM rates; the SA-TG VLM rate (uniform 29-year window) is used to account for VLM influences at tide gauge locations. Therefore, we do not consider additional VLM errors arising from mixing of GNSS VLM rate solutions, as the high-quality constant long-term GNSS VLM estimate is still limited (evident in Lines 242 to 246), so a full replacement of SA-TG VLM rates with GNSS VLM rates for tide gauge VLM adjustment and then re-running the global sea level reconstruction experiment may not be meaningful under current GNSS availability in the global perspective.

In accounting for VLM-related uncertainty in our final reconstruction, we follow prior works: Church & White (2011) and Dangendorf et al. (2019) include the uncertainty associated with the applied VLM correction, and do not consider the VLM errors from the method that we are not involved in. We account for the uncertainty estimation associated with VLM corrections at tide-gauge sites by incorporating the formal error of SA-TG VLM rates into the final reconstructed sea level through 300 Monte Carlo simulation reconstruction (detailed in Lines 413–425). As a result, the uncertainty of our reconstructed sea level at each time step includes the contribution from SA-TG VLM estimates at tide gauge locations, and this contribution is further propagated into the final GMSL trend uncertainty (Lines 453–455).

Finally, we additionally provide the corresponding mean values and more detailed information on Lines 257–263 and edits for easier understanding per reviewers' request.

Detailed changes in the manuscript:

- Lines 257–263 change into: At 253 locations where GNSS and SA-TG are collocated (Figure 3b), the two types of VLM rate estimates show a median (absolute) difference of 0.88 mm yr^{-1} and correlate strongly with $r = 0.86$. The mean VLM rates estimated from 253 locations show 0.50 vs 0.55 mm yr^{-1} from GNSS and SA-TG approaches, respectively. Further analyses were conducted with the 152 GNSS solutions exhibiting the smallest uncertainties among the 253 GNSS-based solutions, with corresponding VLM rates shown in Figure 3c. This subset exhibited a median difference of 0.81 mm yr^{-1} and $r = 0.89$ between GNSS and SA-TG VLM (mean rate: 0.43 vs 0.61 mm yr^{-1}) results. We also examined the 77 GNSS at tide gauge solutions with the smallest uncertainties. The GNSS and SA-TG VLM estimates (mean rate: 0.96 vs 1.11 mm yr^{-1}) show a median difference of 0.64 mm yr^{-1} and a $r=0.95$ (Figure 3d).

RC1-Q12: Lines 283–285: The procedure for transforming the tide gauge time series onto the reference frame of the nearby altimetry point is not clearly explained. Could the authors provide more detail on how this transformation is performed?

Reply: Thank you for the comment. We clarify how the transformation is implemented. At each tide gauge, we first apply the inverse barometer (IB) correction. We then estimate the site VLM as the linear-trend difference between monthly altimetry at the nearest grid point and the IB-corrected tide gauge over their common data span (Jan 1993–Dec 2021). The resulting constant rate is applied to the full 72-year tide-gauge record to express it in the geocentric frame.

Detailed changes in the manuscript:

- For more clarity, we will add one more sentence at Line 283: We estimate the VLM rate for each 72-year tide gauge record as the linear trend difference between the IB-corrected gauge records and the nearest ($\leq 100 \text{ km}$) altimetry sea level measurements over January 1993–December 2021. We then add this constant VLM rate to the entire tide-gauge record to express it in the geocentric (altimetry) reference frame.

RC1-Q13: To improve clarity and reproducibility, I suggest the authors include a schematic flowchart of the data processing workflow. This would provide a more intuitive and comprehensive overview than text descriptions alone.

Reply: We appreciate the suggestion and included a schematic flowchart of the data-processing workflow in the Supplementary Information to help readers quickly grasp the entire processing pipeline.

Detailed changes in the manuscript:

- The flowchart is as follows, and added into the supplementary figure as Fig. S 6:

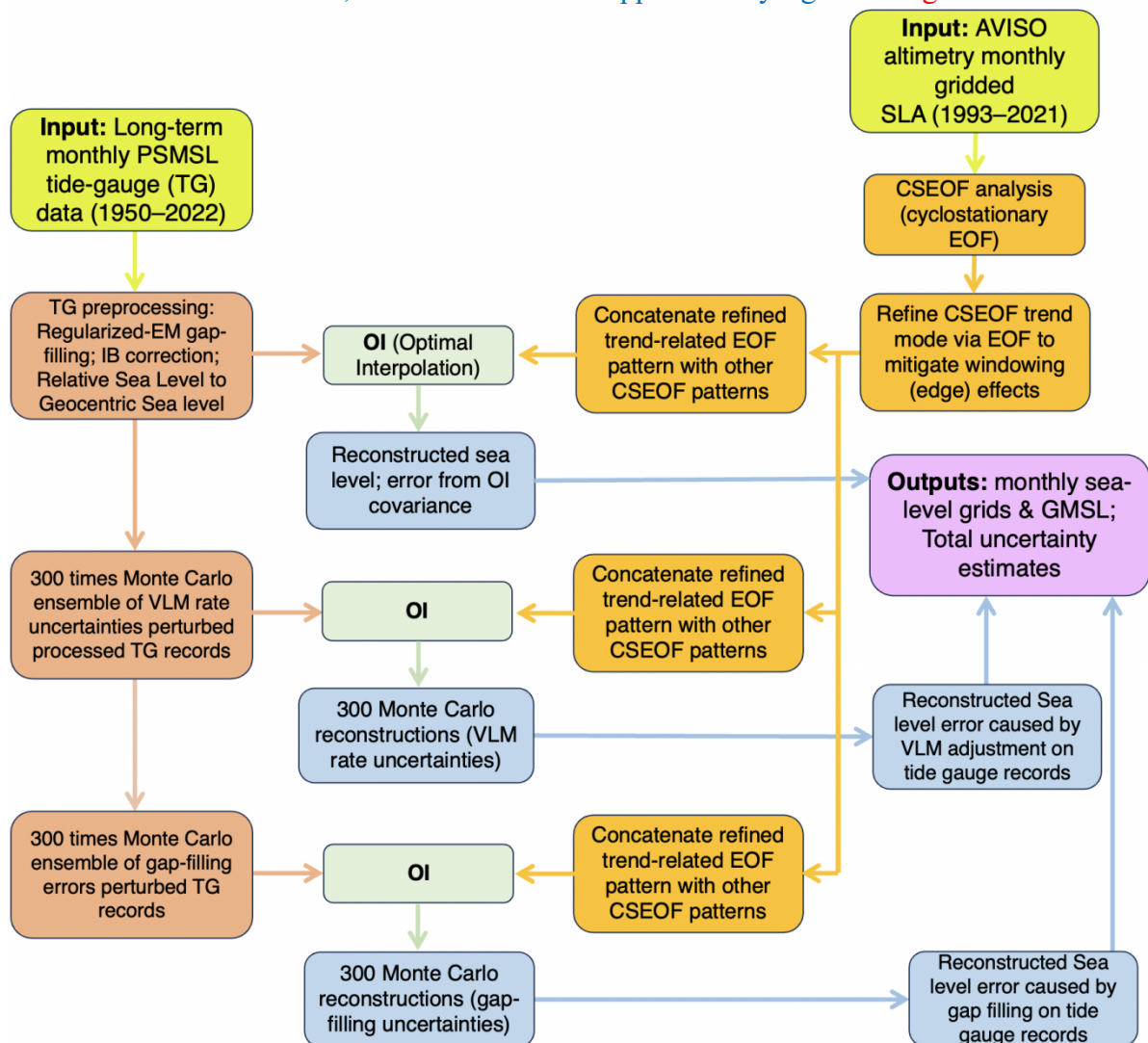


Fig.S 6 (Newly added figure in manuscript). Schematic flowchart of the modified sea-level reconstruction procedure. Inputs are monthly PSMSL tide-gauge records (January 1950–January 2022) and AVISO gridded sea-level anomalies (SLA, January 1993–December 2021). Outputs are 72 years (January 1950–January 2022) of monthly sea-level grids and a global mean sea-level (GMSL) time series, each with corresponding error estimates.

- At line 433, we added a sentence directing readers to the schematic flowchart for a clearer overview of the sea-level reconstruction workflow: For an overview of the modified reconstruction, see the schematic flowchart in Supplementary Information Fig. S6.

RC1-Q14: The manuscript describes the reconstruction of sea level trends from 1950 to 1993 by combining tide gauge and altimetry-derived rates. However, it is unclear how the fusion between tide

gauge records and satellite altimetry data is achieved, given that the AVISO dataset begins in 1993. Specifically, how is the $1^\circ \times 1^\circ$ gridded product generated for the pre-altimetry period, and how are data gaps filled in offshore or deep-ocean regions where tide gauges are absent or sparse?

Reply: Our fusion follows the Kaplan et al. (2000) optimal interpolation framework. The procedure is summarized as follows:

(i) Gridded 29 years (January 1993 to December 2021) monthly satellite-altimetry sea-level anomalies compute the leading spatial pattern (in this study: 1 EOF pattern for accounting for trend plus 19 CSEOF patterns); (ii) For the full reconstruction period (Jan 1950–Jan 2022, 865 months), we regress these patterns onto tide-gauge anomalies (VLM- and IB-corrected) at tide-gauge locations to obtain monthly time coefficients (A_R in equation 11). This yields 865 coefficients per pattern. (iii) Multiplying each pattern by its monthly coefficient, we can generate one mode. In our case, there are 20 patterns, so we have 20 modes, and summing across the 20 modes gives monthly $1^\circ \times 1^\circ$ sea-level fields spanning 865 months (from January 1950 to January 2022, including the pre-altimetry period).

Because our ability to determine where we can hold reconstruction depends on where the altimetry-derived spatial patterns are defined, regions covered by those patterns—no matter if coastal, offshore, or in the open ocean—are included in the reconstruction. The optimal-interpolation step provides formal uncertainties via the error variance–covariance matrix (Equation 10). Beyond these optimal interpolation process errors, we additionally quantify the reconstructed sea level uncertainties from tide-gauge VLM estimation and from tide-gauge gap filling as mentioned in Lines 416–425, section 2.3.

RC1-Q15: The manuscript refers to the application of a Reduced Space Optimal Interpolation method based on CSEOFs (CSEOF-OI), but the underlying algorithm is not clearly described. For reproducibility and clarity, I suggest the authors provide a concise explanation of how the CSEOF decomposition and the subsequent optimal interpolation are implemented, and how this approach improves upon traditional EOF-OI methods. A schematic or reference to a methodological appendix would be helpful.

Reply: Thank you for the suggestion. Our manuscript focuses on an improved CSEOF-OI implementation rather than rederiving the original CSEOF-OI algorithm. To support reproducibility without duplicating prior work, we (i) added a one-sentence pointer in Section 2.3 to the CSEOF formulations and tutorials, and (ii) appended a brief implementation note after Equations. (8)-(12) summarizing the reduced-space OI steps (how R , Q , and A_R how $\hat{x}$ is reconstructed). We have clarified how our approach improves on contemporary EOF-OI by mitigating the CSEOF trend "windowing/edge" effect through an additional EOF refinement of the trend mode, which reduces propagated heteroscedasticity in the extended period (see Figure 4, Figure 5, and related text).

Detailed changes in the manuscript:

- Add one sentence in Line 329: For algorithmic details and practical guidance on CSEOF, we direct readers to Kim and North (1997) and Kim (2015) for tutorials and applications; applications to CSEOF-OI sea-level reconstruction can be found in Hamlington et al. (2011) and Hamlington et al. (2014).
- Before Line 381, we will add: For completeness, we include a brief implementation note describing the OI steps; detailed derivations follow Kaplan et al. (2000).

RC1-Q16: In Figure 6, panels (a) and (b) appear to show very similar results, despite panel (b) involving additional processing steps such as EOF or CSEOF decomposition. Since both panels are based on the same altimetry data, this comparison mainly reflects differences introduced by the processing itself rather than demonstrating any improvement in the quality of the reconstructed data. I suggest that the authors clarify the purpose of this comparison and provide more rigorous evidence to support claims of improved data quality.

Reply: Thank you for the comment. The comparison between Figure 6 panels (a) and (b) is a standard validation used to assess reconstructed sea-level quality during the satellite-altimetry era; similar comparisons appear in Hamlington et al. (2011), Dangendorf et al. (2024), and Wang et al. (2024). Note that panel (b) shows the linear trend fitted to the reconstructed sea-level time series during the altimetry era, rather than results taken from spatial-temporal decomposition. The purpose of Figure 6a–b is not to claim that (b) improves upon (a) but to provide an altimetry-era consistency check: when both fields are evaluated over January 1993–December 2021, our reconstruction can well reproduce the regional trends from measurements, so the visual similarity is expected.

Detailed changes in the manuscript:

- For better understanding, we add one more explanation at Line 447 into: (c) illustrates their difference; for altimetry-era trend consistency check, and the close similarity between (a) and (b) indicates that the reconstruction efficiently reproduces the altimetry-observed trend field.

Note: Q17 is related to Q11, and its question and related response are already shown right after Q10.

RC1-Q18: The reconstructed dataset is provided at a $1^\circ \times 1^\circ$ resolution; however, the validation analyses are primarily conducted at $5^\circ \times 5^\circ$ or coarser spatial scales. It is important to assess the accuracy and stability of the reconstruction at the native grid resolution, especially for users interested in regional-scale applications. Furthermore, the quality assessment largely focuses on trend estimates and correlations with known climate indices. While useful, these do not sufficiently address the reliability of the reconstructed sea level fields in the pre-altimetry era (before 1993), particularly in the open ocean where tide gauge constraints are absent. Greater emphasis should be placed on evaluating the uncertainty and credibility of the reconstructed fields in such data-sparse regions.

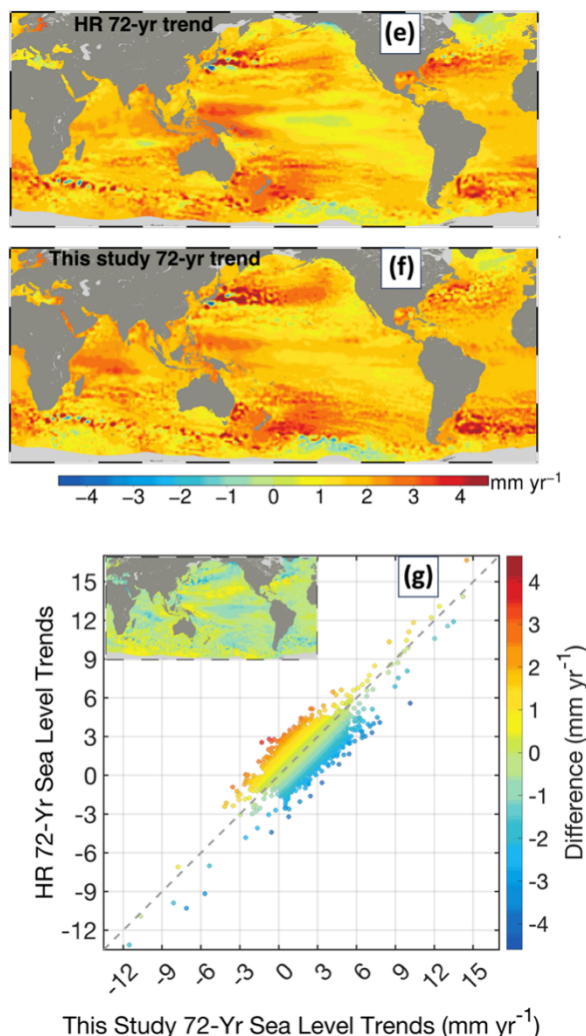
Reply: We agree that users benefit from validation at the native $1^\circ \times 1^\circ$ grid and from a clearer assessment of reliability in the pre-altimetry era, especially over the open ocean. To achieve this, we will:

(i) Revise Figure 6 panels e, f, and g to present $1^\circ \times 1^\circ$ comparisons to evaluate the 72-year regional sea level trend.

(ii) For the reliability of the reconstructed sea-level time series in the pre-altimetry era (before 1993), we validated the results using monthly long-term (40 years and 65 years) tide gauges that were not used in the reconstruction; this is already shown in Figure 6d. For the open ocean prior to 1993, our algorithm provides an uncertainty estimate at every time step for reconstructed time series at each $1^\circ \times 1^\circ$ grid. Because this component was not displayed in the manuscript figures, it was not included in our initial data release. We have now further uploaded the uncertainty (“error”) data and respective user guidance in the updated data availability link:

<https://doi.org/10.5281/zenodo.15288816> for the original dataset submitted. And we selected several representative open-ocean locations for visualization here (illustrated in RC1_Fig2).

Detailed changes in the manuscript:



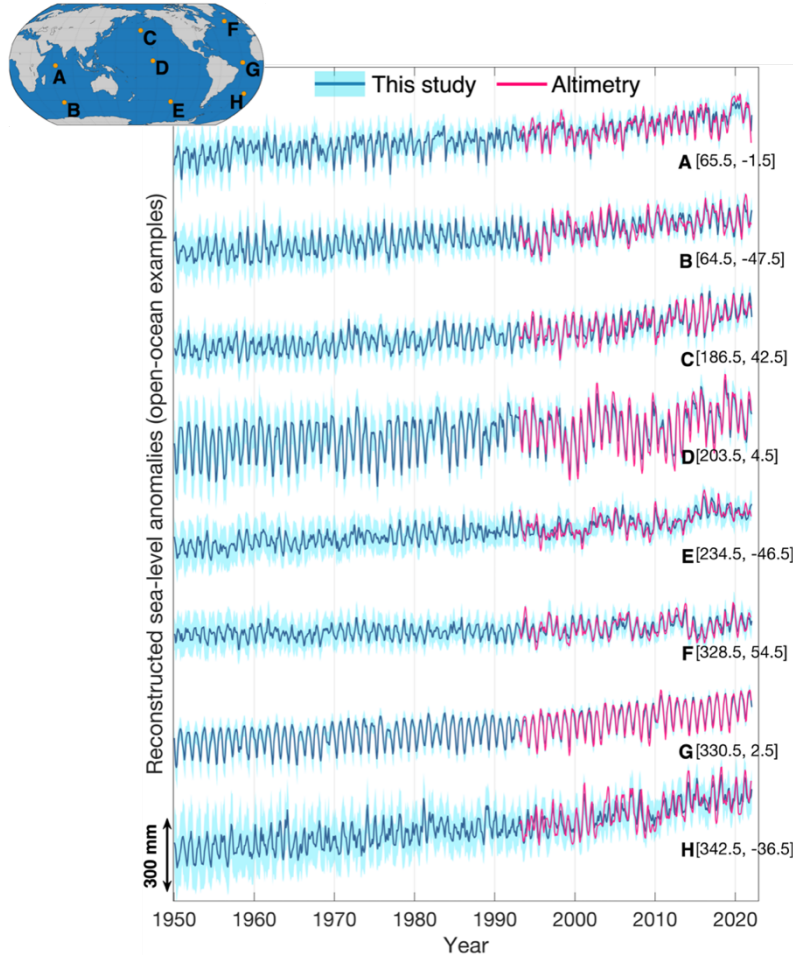
- Revise the context from Lines 470 to 475 into: Due to the spatial resolution difference between the two reconstruction products, we examine sea level trends from 1950 to 2021 across each $1^\circ \times 1^\circ$ grid in various ocean basins by first resampling the products from Dangendorf et al. (2024) with the exact spatial resolution with our product, then hold the comparison across each grid (Within our 1° ocean grid, >99.9% of cells have a corresponding value in the re-gridded Dangendorf field), as detailed in Figure 6e- f. Further analysis based on Figure 6g reveals that ~40% of the $1^\circ \times 1^\circ$ grids exhibit discrepancies (absolute) below 0.30 mm yr^{-1} , while ~80% of the grids show discrepancies under 0.74 mm yr^{-1} , and ~92% maintain discrepancies less than 1 mm yr^{-1} .

- Figure 6 (e,f,g, revised): Panels (e)–(f) depict regional geocentric sea-level trend over 1950–2021 from HR (Dangendorf et al. 2024, re-gridded at $1^\circ \times 1^\circ$) and from the modified reconstruction, (g) highlights the regional differences.

- And we will add a small paragraph after line 425 for more details about the regional error estimate: We denote the per-grid uncertainty contributions from gap filling and from the VLM correction at time t as $\sigma_{\hat{x}_{filling}}(i, t)$ and $\sigma_{\hat{x}_{VLM}}(i, t)$, respectively, both are defined as the one standard deviation spread across the Monte-Carlo reconstructions in which the corresponding component is perturbed and the full reconstruction is rerun. Together with the error from the optimal interpolation process. The uncertainty for the i^{th} grid at time t can be represented as:

$$\sigma_{\text{regional}}(i, t) \approx \sqrt{\text{Cov}H(i, i, t) + \left(\sigma_{\hat{x}_{filling}}(i, t)\right)^2 + \left(\sigma_{\hat{x}_{VLM}}(i, t)\right)^2} \quad (16)$$

Visualization of several representative reconstructed sea levels in open-ocean locations regarding the reviewer's concern:



RC1_Fig2. Open-ocean examples of reconstructed sea-level anomalies (cyan) with satellite-altimetry anomalies (magenta) at eight locations (A–H). Shaded envelopes show the reconstruction uncertainty as $\pm 3\sigma$ (three standard deviations). The longitude–latitude coordinates [Lon (0 to 360), Lat] are listed to the right of each timeseries, and the corresponding sites are marked on the inset map (upper left).

References:

- Bevis, M., Scherer, W., and Merrifield, M.: Technical Issues and Recommendations Related to the Installation of Continuous GPS Stations at Tide Gauges, *Marine Geodesy*, 25, 87–99, <https://doi.org/10.1080/014904102753516750>, 2002.
- Bruni, S., Fenoglio, L., Raicich, F., and Zerbini, S.: On the consistency of coastal sea-level measurements in the Mediterranean Sea from tide gauges and satellite radar altimetry, *J Geod*, 96, 41, <https://doi.org/10.1007/s00190-022-01626-9>, 2022.
- Calafat, F. M., Chambers, D. P., and Tsimplis, M. N.: On the ability of global sea level reconstructions to determine trends and variability, *J. Geophys. Res. Oceans*, 119, 1572–1592, <https://doi.org/10.1002/2013JC009298>, 2014.
- Church, J. A. and White, N. J.: Sea-Level Rise from the Late 19th to the Early 21st Century, *Surv Geophys*, 32, 585–602, <https://doi.org/10.1007/s10712-011-9119-1>, 2011.
- Dangendorf, S., Marcos, M., Wöppelmann, G., Conrad, C. P., Frederikse, T., and Riva, R.: Reassessment of 20th century global mean sea level rise, *Proc. Natl. Acad. Sci. U.S.A.*, 114, 5946–5951, <https://doi.org/10.1073/pnas.1616007114>, 2017.
- Dangendorf, S., Hay, C., Calafat, F. M., Marcos, M., Piecuch, C. G., Berk, K., and Jensen, J.: Persistent acceleration in global sea-level rise since the 1960s, *Nat. Clim. Chang.*, 9, 705–710, <https://doi.org/10.1038/s41558-019-0531-8>, 2019.
- Dangendorf, S., Sun, Q., Wahl, T., Thompson, P., Mitrovica, J. X., and Hamlington, B.: Probabilistic reconstruction of sea-level changes and their causes since 1900, *Earth Syst. Sci. Data*, 16, 3471–3494, <https://doi.org/10.5194/essd-16-3471-2024>, 2024.
- Ebisuzaki, W.: A Method to Estimate the Statistical Significance of a Correlation When the Data Are Serially Correlated, *J. Climate*, 10, 2147–2153, [https://doi.org/10.1175/1520-0442\(1997\)010<2147:AMTETS>2.0.CO;2](https://doi.org/10.1175/1520-0442(1997)010<2147:AMTETS>2.0.CO;2), 1997.
- Hamlington, B. D., Leben, R. R., Nerem, R. S., Han, W., and Kim, K.-Y.: Reconstructing sea level using cyclostationary empirical orthogonal functions, *J. Geophys. Res.*, 116, C12015, <https://doi.org/10.1029/2011JC007529>, 2011.
- Hamlington, B. D., Leben, R. R., Strassburg, M. W., and Kim, K. -Y.: Cyclostationary empirical orthogonal function sea-level reconstruction, *Geosci. Data J.*, 1, 13–19, <https://doi.org/10.1002/gdj3.6>, 2014.
- Hammond, W. C., Blewitt, G., Kreemer, C., and Nerem, R. S.: GPS Imaging of Global Vertical Land Motion for Studies of Sea Level Rise, *JGR Solid Earth*, 126, e2021JB022355, <https://doi.org/10.1029/2021JB022355>, 2021.
- He, X., Huang, J., Montillet, J.-P., Wang, S., Kermarrec, G., Shum, C. K., Hu, S., and Wang, F.: A Noise Reduction Approach for Improve North American Regional Sea Level Change from Satellite and In Situ Observations, *Surv Geophys*, <https://doi.org/10.1007/s10712-025-09894-8>, 2025.

Kaplan, A., Kushnir, Y., and Cane, M. A.: Reduced Space Optimal Interpolation of Historical Marine Sea Level Pressure: 1854–1992*, *J. Climate*, 13, 2987–3002, [https://doi.org/10.1175/1520-0442\(2000\)013<2987:RSOIOH>2.0.CO;2](https://doi.org/10.1175/1520-0442(2000)013<2987:RSOIOH>2.0.CO;2), 2000.

Kim, K.-Y. and North, G. R.: EOFs of Harmonizable Cyclostationary Processes, *J. Atmos. Sci.*, 54, 2416–2427, [https://doi.org/10.1175/1520-0469\(1997\)054<2416:EOHCP>2.0.CO;2](https://doi.org/10.1175/1520-0469(1997)054<2416:EOHCP>2.0.CO;2), 1997.

Kim, K.-Y., Hamlington, B., and Na, H.: Theoretical foundation of cyclostationary EOF analysis for geophysical and climatic variables: Concepts and examples, *Earth-Science Reviews*, 150, 201–218, <https://doi.org/10.1016/j.earscirev.2015.06.003>, 2015.

Kuo, C. Y.: Vertical crustal motion determined by satellite altimetry and tide gauge data in Fennoscandia, *Geophys. Res. Lett.*, 31, L01608, <https://doi.org/10.1029/2003GL019106>, 2004.

Kuo, C.-Y., Shum, C. K., Braun, A., Cheng, K.-C., and Yi, Y.: Vertical Motion Determined Using Satellite Altimetry and Tide Gauges, *Terr. Atmos. Ocean. Sci.*, 19, 21, [https://doi.org/10.3319/TAO.2008.19.1-2.21\(SA\)](https://doi.org/10.3319/TAO.2008.19.1-2.21(SA)), 2008.

Mann, M. E., Rutherford, S., Wahl, E., and Ammann, C.: Robustness of proxy-based climate field reconstruction methods, *J. Geophys. Res.*, 112, 2006JD008272, <https://doi.org/10.1029/2006JD008272>, 2007.

Nerem, R. S. and Mitchum, G. T.: Estimates of vertical crustal motion derived from differences of TOPEX/POSEIDON and tide gauge sea level measurements, *Geophysical Research Letters*, 29, <https://doi.org/10.1029/2002GL015037>, 2002.

Oelsmann, J., Passaro, M., Dettmering, D., Schwatke, C., Sánchez, L., and Seitz, F.: The zone of influence: matching sea level variability from coastal altimetry and tide gauges for vertical land motion estimation, *Ocean Sci.*, 17, 35–57, <https://doi.org/10.5194/os-17-35-2021>, 2021.

Peltier, W. R.: GLOBAL GLACIAL ISOSTASY AND THE SURFACE OF THE ICE-AGE EARTH: The ICE-5G (VM2) Model and GRACE, *Annu. Rev. Earth Planet. Sci.*, 32, 111–149, <https://doi.org/10.1146/annurev.earth.32.082503.144359>, 2004.

Ray, R. D., Beckley, B. D., and Lemoine, F. G.: Vertical crustal motion derived from satellite altimetry and tide gauges, and comparisons with DORIS measurements, *Advances in Space Research*, 45, 1510–1522, <https://doi.org/10.1016/j.asr.2010.02.020>, 2010.

Santamaría-Gómez, A. and Mémin, A.: Geodetic secular velocity errors due to interannual surface loading deformation, *Geophysical Journal International*, 202, 763–767, <https://doi.org/10.1093/gji/ggv190>, 2015.

Santamaría-Gómez, A., Gravelle, M., Collilieux, X., Guichard, M., Míguez, B. M., Tiphaneau, P., and Wöppelmann, G.: Mitigating the effects of vertical land motion in tide gauge records using a state-of-the-art GPS velocity field, *Global and Planetary Change*, 98–99, 6–17, <https://doi.org/10.1016/j.gloplacha.2012.07.007>, 2012.

Santamaría-Gómez, A., Gravelle, M., Dangendorf, S., Marcos, M., Spada, G., and Wöppelmann, G.: Uncertainty of the 20th century sea-level rise due to vertical land motion errors, *Earth and Planetary Science Letters*, 473, 24–32, <https://doi.org/10.1016/j.epsl.2017.05.038>, 2017.

Steig, E. J., Schneider, D. P., Rutherford, S. D., Mann, M. E., Comiso, J. C., and Shindell, D. T.: Warming of the Antarctic ice-sheet surface since the 1957 International Geophysical Year, *Nature*, 457, 459–462, <https://doi.org/10.1038/nature07669>, 2009.

Schneider, T.: Analysis of Incomplete Climate Data: Estimation of Mean Values and Covariance Matrices and Imputation of Missing Values, *J. Climate*, 14, 853–871, [https://doi.org/10.1175/1520-0442\(2001\)014<0853:AOICDE>2.0.CO;2](https://doi.org/10.1175/1520-0442(2001)014<0853:AOICDE>2.0.CO;2), 2001.

Wan, J.: Joint estimation of vertical land motion and global sea-level rise over the past six decades using satellite altimetry and tide gauge records, PhD thesis, The Ohio State University, Columbus, OH, USA, http://rave.ohiolink.edu/etdc/view?acc_num=osu1449185593, 2015.

Wang, J., Church, J. A., Zhang, X., and Chen, X.: Improved Sea Level Reconstruction from 1900 to 2019, *Journal of Climate*, 37, 6453–6474, <https://doi.org/10.1175/JCLI-D-23-0410.1>, 2024.

Williams, S. D. P., Bock, Y., Fang, P., Jamason, P., Nikolaidis, R. M., Prawirodirdjo, L., Miller, M., and Johnson, D. J.: Error analysis of continuous GPS position time series, *J. Geophys. Res.*, 109, 2003JB002741, <https://doi.org/10.1029/2003JB002741>, 2004.

Wöppelmann, G. and Marcos, M.: Vertical land motion as a key to understanding sea level change and variability, *Reviews of Geophysics*, 54, 64–92, <https://doi.org/10.1002/2015RG000502>, 2016.

Wöppelmann, G., Gravelle, M., Guichard, M., and Prouteau, E.: Progress report on the GNSS at tide gauge data assembly center: SONEL data holdings & tools to access the data, status report, GLOSS-GE Meeting, Busan, Republic of Korea, 11–13 April 2019, available at: <https://www.sonel.org>, 2019.

Wöppelmann, G., Letetrel, C., Santamaria, A., Bouin, M. -N., Collilieux, X., Altamimi, Z., Williams, S. D. P., and Miguez, B. M.: Rates of sea-level change over the past century in a geocentric reference frame, *Geophysical Research Letters*, 36, 2009GL038720, <https://doi.org/10.1029/2009GL038720>, 2009.