Response to Reviewer #2 Comments

Comment 1:

This paper models the effect of the thermoelastic component of ground displacement driven by ERA5 temperatures down to 2 m of depth and soil properties also down to 2 m using a soil data base. It presents the result of a thermoelastic model at > 15,000 sites and it provides time series for the model output for each station in a companion website. It is evident that the sub millimeter annual amplitudes of the displacements for the layered model are different than those for a half-space model as the physics would predict and might warrant inclusion into GNSS post-processing analyses. The paper represents a tremendous amount of research.

90% of the manuscript is dedicated to comparing the half space model with the layered model. It is clear that the physics of the layered model represents an improvement over the half-space model. But this is not tested with real observations. At the end of the paper, there is a comparison of the layered model with GNSS observations. The improvement of the model to the WRMS of the GNSS data is very minor. But this is where the half-space model and the layered model should be compared. The WRMS reduction of the GNSS observations using both models should be compared.

The paper is generally well written and well referenced. It could be shortened substantially that would help to make it more focused. There are just too many comparisons of the half space model with the layered model. What the reader wants to know is if the correction reduces the RMS of the GNSS observations in any substantial way.

Response 1:

Thank you very much for your positive assessment of the dataset and for the constructive suggestions. We agree that the manuscript should be more focused on the released TED product and its practical relevance to GNSS height time series, rather than placing too much emphasis on repeated comparisons between the homogeneous and layered models. In the revised manuscript, we have reorganized and shortened the Results and Discussion sections to make this focus clearer.

Specifically, we reduced the discussion centred on the homogeneous–layered model comparison and retained it only where it helps quantify the effect of subsurface stratification. We also revised Section 3.3 so that the GNSS assessment now begins with the TED-only WRMS change. In addition, we added a direct comparison of the WRMS changes obtained from the FEMFL correction and the traditional homogeneous FEM correction, so that the two models are evaluated against real GNSS observations. The Abstract and Discussion and conclusions were also revised to avoid overstating the combined correction with non-tidal loading products and to report the standalone contribution of TED more clearly.

Overall, the revised manuscript now presents the homogeneous–layered model comparison as supporting evidence for the dataset construction, while the main practical assessment focuses on the TED correction applied to GNSS height time series.

Comment 2: The biggest concern is a robust validation of the model. Does the RMS of the GNSS up time series decrease across the majority of stations when the model is applied? The reader does not learn the answer to this question until page 22-23. The abstract states 'When TED corrections are applied together with non-tidal mass loading correction, the residual vertical dispersion decreases at most stations, with vertical scatter reduced by up to 70% at selected sites'. This analysis does not test the thermoelastic model in isolation. It tests the thermoelastic model + large amplitude loading effects. We cannot determine whether the TED corrections or the loading models cause the change in the RMS.

Response 2:

Thank you for this important comment. We agree that the evaluation of the TED product should be based on the thermoelastic correction itself, rather than on the combined correction with non-tidal loading products. In the revised manuscript, we removed the statement in the Abstract that reported the WRMS reduction from the combined TED + loading correction. The Abstract now reports only the TED-only correction result, stating that applying TED alone reduces WRMS at many stations, with the largest WRMS reductions reaching about 10%.

We also reorganized Section 3.3 so that the GNSS-based assessment now begins with the TED-only WRMS change. In this assessment, only the TED correction is applied to the GNSS height time series, without removing NTAL, NTOL, or HYDL. The revised text makes clear that the TED-only effect is modest and spatially variable, mostly within 3–5% and locally reaching about 5–10%, while some stations show increased WRMS after correction. The loading products are now used only as background comparisons for signal magnitude and timing, and their combined effect is no longer used to demonstrate the performance of the TED product.

Revised text in the Abstract:

"When applied alone to GNSS height time series, TED produces positive WRMS reductions at about 79% of the assessed stations, with most improvements within 3–5% and the largest WRMS reductions reaching about 10%."

Revised text in Section 3.3:

"The effect of TED on GNSS height time series is first examined by applying the TED correction alone, without removing NTAL, NTOL, or HYDL (Fig. 11). Positive WRMS reductions are found at about 79% of the assessed stations, indicating that removing TED reduces the scatter of GNSS height time series at most sites. The improvement is generally modest, mostly within 3–5%, and locally reaches about 5–10%. Meanwhile, some stations show negative WRMS reductions, locally around −2% to −5%, indicating that the TED-only correction does not improve every GNSS height time series."

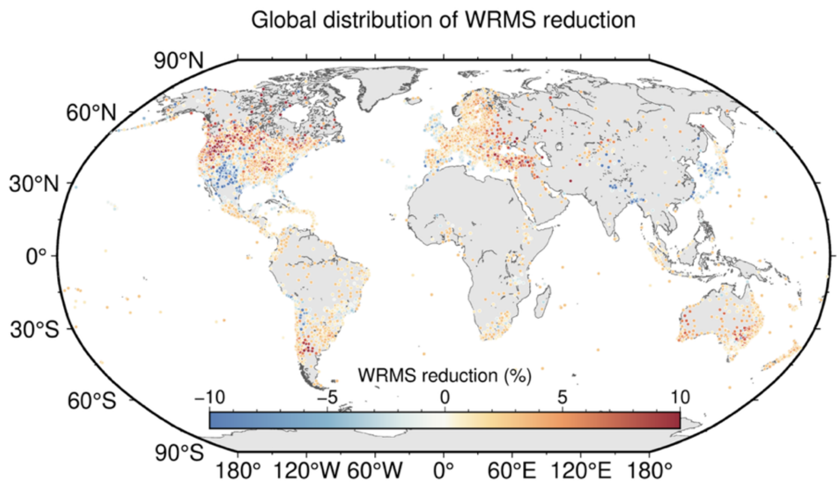

Figure 11 Global distribution of WRMS changes after applying TED-only correction to GNSS height time series. Positive values indicate reduced WRMS after removing TED, while negative values indicate increased WRMS.

Comment 3: The paper does not present any order of magnitude comparisons for the reader. How big are these Earth thermoelastic effects compared to monument thermoelastic effects? How big are these thermoelastic effects with respect to loading (there is a figure late in the paper but a figure showing the global effects of a half-space model versus the loading early in the paper would be useful. And the figure should be a global figure not only a few stations.) How sensitive is the model to sediment thicknesses and thermal/mechanical properties? In this vein, it would be useful to see a plot of sediment thicknesses over the continents. When I compare the differences between the half space model and the Global 1-km Gridded Thickness of Soil, Regolith, and Sedimentary Deposit Layers from ORNL DAAC there is a lot of correlation.

Response 3:

Thank you for this helpful suggestion. We agree that the manuscript should provide clearer order-of-magnitude comparisons so that readers can evaluate the practical significance of the TED product. In the revised manuscript, we added a global order-of-magnitude comparison in Section 3.1, where the annual amplitudes of the FEMFL TED product are compared with the traditional homogeneous FEM solution, above-ground monument thermal expansion, and non-tidal loading displacements, including NTAL, NTOL, and HYDL. This comparison is now presented early in the Results section rather than only through a few representative stations.

We have also strengthened the discussion of model sensitivity in Section 3.2. The revised manuscript now summarizes the response of the annual TED cycle to perturbations in the main thermal and mechanical parameters, including thermal expansion coefficient, thermal diffusivity, and Young's modulus. These results show that the annual amplitude and phase are generally stable under reasonable parameter uncertainty, although larger sensitivity can occur in regions with stronger near-surface thermal contrasts. In addition, we clarified that shallow subsurface structure is one reason for the spatially coherent differences between the layered and homogeneous models.

Regarding sediment or regolith thickness, we agree that this is an important factor for interpreting the difference between the layered and homogeneous models. In the current release, the layered structure is constructed from SoilGrids in the upper 2 m, and deeper sediment-thickness information is not used as a direct model input. We have therefore clarified this point in the revised manuscript and discussed sediment/regolith thickness as a likely contributor to the spatial pattern of layered–homogeneous differences. A more complete integration of continental sediment-thickness datasets, such as the ORNL DAAC product, will be considered in future versions of the model.

Revised text added in Section 3.1:

"Figure 5 compares the annual amplitude of TED with monument thermal expansion and non-tidal loading displacements. The median annual amplitude is about 0.50 mm for the traditional homogeneous FEM solution and about 0.71 mm for the FEMFL product, indicating that introducing depth-dependent subsurface properties generally increases the annual TED amplitude. The FEMFL TED amplitude is comparable to monument thermal expansion, whose median value is about 0.40 mm for stations with available monument information. Compared with non-tidal loading, TED is generally smaller than hydrological loading, but its annual amplitude overlaps with the lower-to-middle ranges of NTOL and NTAL. These comparisons suggest that TED is a small but measurable correction component for high-precision GNSS height time series."

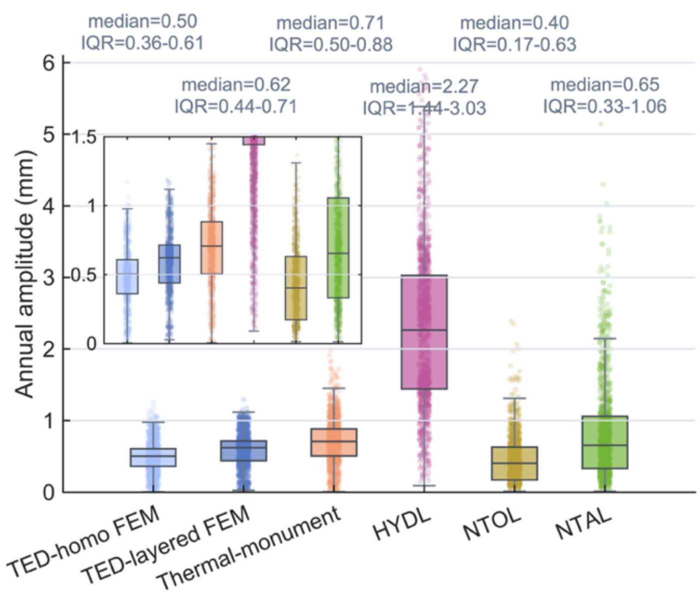

Figure 5 Order-of-magnitude comparison of annual vertical displacement amplitudes from TED, monument thermal expansion, and non-tidal loading at 2,612 GNSS stations with documented monument information. TED-FEM denotes the traditional homogeneous FEM solution, while TED-FEM_{FL} denotes the FEM_{FL} solution used for the released TED product. Thermal-monument represents above-ground monument thermal expansion estimated from the available monument information. HYDL, NTOL, and NTAL denote hydrological, non-tidal oceanic, and non-tidal atmospheric loading displacements,

respectively. The inset enlarges the TED and monument-thermal-expansion amplitudes for comparison.

Revised text added in Section 3.2:

"The parameter-sensitivity results are summarized in Fig. 8. The linear thermal expansion coefficient α, thermal diffusivity κ, and Young's modulus E are each perturbed by ±10%, and the resulting changes in the annual TED amplitude and phase are examined. The amplitude response is generally limited. The largest sensitivity occurs mainly in dry inland regions, where stronger seasonal temperature forcing also produces larger TED signals. In these regions, most annual-amplitude changes remain below about 0.15 mm, although a small number of stations show larger responses. Humid coastal regions usually show weaker sensitivity, mostly below 0.1 mm, while high-latitude regions show intermediate behaviour. Among the three tested parameters, α and κ have a more direct influence on the annual amplitude, whereas the effect of E is generally smaller. The phase response is also moderate, with most annual-phase changes within about 1–3° and only a limited number of stations reaching several degrees."

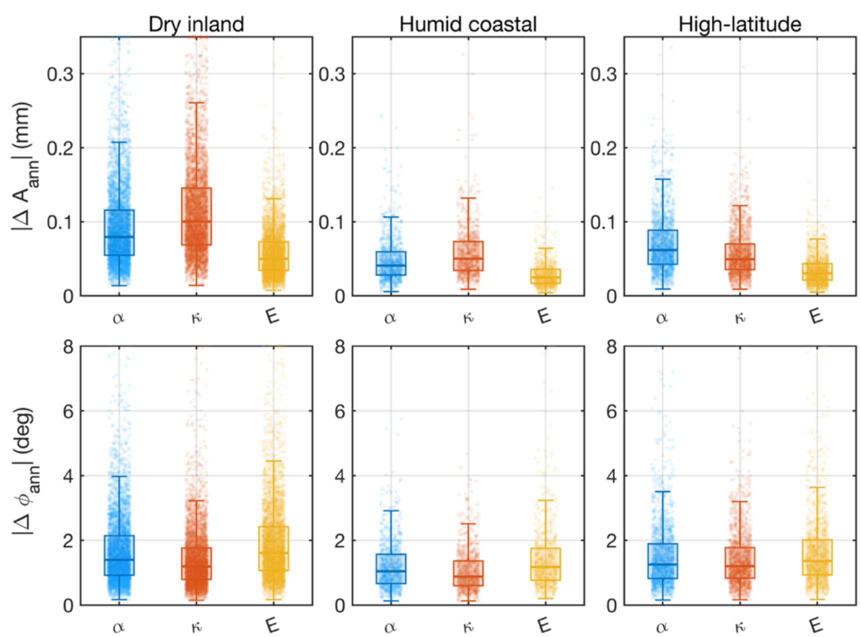

Figure 8 Distribution of TED annual-cycle sensitivity under ±10% parameter perturbations. Box-and-scatter summaries for three region types (arid interior, humid coastal, and high latitude), showing changes in the TED annual cycle when α (linear thermal expansion coefficient), κ (thermal diffusivity), and E (Young's modulus) are perturbed by ±10%.

Minor comments

Comment 4: Soil model: The authors use sediment characteristics down to 2 m. For greater depths they assume a half-space model assumed to be bedrock. Do you stop at 2 m because the thermoelastic effect diminishes at this depth or because you do not have model data.

1. It would be useful if they presented the results for an order of magnitude calculation. To what depth are thermoelastic effects important? Of course this varies by material. But the authors could provide a half-space model for two end points of high thermally conductive and low thermally conductive material to demonstrate how deep the effect can go.
2. ALSO, what does the soil model look like for the edges of ice covered regions where the temperature range is usually damped by the low thermal conductivity of the oceans?
3. Can you include a representative figure of your soil model?
4. Can you provide a description of the half-space model? Sorry if I missed this. but I've gone through the paper twice trying to find this description.
5. It appears soil moisture content of the soils are not considered. Could you provide an amplitude analysis and a discussion of how the results would change with changes in soil moisture content? The authors mention vegetation as a controlling factor. It isn't the vegetation; it is most likely the soil moisture.

Response 4:

Thank you for these detailed and helpful comments. We agree that the shallow soil model, the treatment below 2 m, and the role of soil moisture need to be described more clearly. In the revised manuscript, we clarified that the use of SoilGrids down to 2 m reflects the depth range of the available global soil-property dataset, rather than an assumption that thermoelastic effects vanish below 2 m. Below 2 m, the model uses a homogeneous bedrock background with prescribed thermal and elastic properties. We have added this description in Sections 2.3 and 2.4.

In addition, we have added an order-of-magnitude discussion of thermal penetration depth. Using the characteristic penetration depth $\delta = \sqrt{\kappa P/\pi}$, the diurnal temperature signal is mainly confined to the upper about 0.1–0.2 m for typical near-surface thermal diffusivities, whereas the annual signal can reach about 1.7–3.2 m. This explains why the upper 2 m layering is important, while also showing that deeper annual-scale thermal effects need to be represented by the simplified homogeneous bedrock background. We have added this calculation to clarify the depth range over which thermoelastic effects are expected to be important.

For ice-covered regions, we agree that the thermal boundary condition and subsurface properties are difficult to represent reliably near permanent ice sheets, glaciers, and ice shelves. In the revised dataset and manuscript, stations located in Antarctica, Greenland, permanent ice-covered or glacier-covered regions, ice shelves, and their surrounding buffer zones have been excluded from the released dataset and the subsequent statistical analyses. This avoids applying the SoilGrids-based land model to regions where the shallow subsurface and thermal boundary conditions are not well represented.

We also agree that soil moisture is a key control on the effective thermal properties of the shallow subsurface. The current model does not explicitly include time-varying soil moisture. We revised the text to avoid attributing this effect to vegetation itself and now describe soil moisture as a likely controlling factor. The sensitivity tests on thermal diffusivity provide a first-order indication of how changes in effective thermal properties can affect annual TED amplitude and phase, but they do not replace a fully soil-moisture-dependent thermal model. We have added this point as a limitation and as an important direction for future model development.

Revised text added in Section 2.3:

"Below 2 m, globally consistent high-resolution soil-property information is not available from SoilGrids. The use of 2 m as the lower limit of the SoilGrids-derived layers therefore reflects the depth range of the available global soil dataset, rather than an assumption that thermoelastic effects vanish below this depth. In the model, the medium below 2 m is represented by a homogeneous bedrock background with prescribed thermal and elastic properties. This treatment preserves the shallow-layer contrasts represented by SoilGrids while using a simplified description of the deeper thermal and elastic background."

Revised text added in Section 2.4:

"The layered column consists of SoilGrids-derived layers in the upper 2 m and a homogeneous bedrock background below. The thermophysical, elastic, and thermoelastic parameters of each layer are assigned using the SoilGrids-based parameterization described in Section 2.3 and are assumed to be uniform within each layer. The deeper background is assigned prescribed thermal and elastic properties and represents the deeper thermal and elastic conditions where globally consistent station-scale subsurface information is not available."

"As an order-of-magnitude estimate, the characteristic thermal penetration depth can be approximated by $\delta = \sqrt{\kappa P/\pi}$, where κ is thermal diffusivity and P is the forcing period. For typical near-surface diffusivities of 0.3×10^{-6} to 1.0×10^{-6} $m^2\ s^{-1}$, diurnal temperature variations mainly affect the upper 0.1–0.2 m, whereas annual temperature variations can reach about 1.7–3.2 m. The SoilGrids-based layering in the upper 2 m therefore captures the main shallow subsurface contrasts, while deeper annual-scale thermal effects are represented in a simplified way by the homogeneous bedrock background."

Revised text added in the limitations:

"The present subsurface parameterization also does not explicitly include time-varying soil moisture. Because soil moisture can modify thermal conductivity, volumetric heat capacity, and effective thermal diffusivity, its influence is treated as a source of uncertainty and is partly reflected in the parameter-sensitivity tests."

Comment 5: P9; L 277: where are these results presented?

Response 5: The original sentence referred to a broader WRMS comparison involving NTAL, NTOL, HYDL, TED, and their combined correction. However, following the reviewer's earlier comments, we have revised the manuscript to focus the GNSS-based

assessment on TED correction alone. Therefore, this sentence has been revised to avoid referring to WRMS results that are no longer presented as part of the main analysis.
Revised:
"where $WRMS_{raw}$ and $WRMS_{corr}$ denote WRMS of the original GNSS height time series and the WRMS after applying TED correction alone, respectively."

Comment 6: Sensitivity analysis: where are the results shown? What are volcano plots?

Response 6: The term "volcano plots" was unclear and not appropriate in this context. We have removed this wording from the revised manuscript.

We have also revised the text to make clear where the sensitivity-analysis results are presented. The sensitivity tests are now described in Section 2.5 and the results are presented in Section 3.2 and Fig. 8. These tests examine the response of the annual TED amplitude and phase to perturbations in key thermal and elastic parameters, including the linear thermal expansion coefficient, thermal diffusivity, and Young's modulus.

Comment 7: Figure 3: where are these stations and what climatological regimes do they represent? Are these daily model solutions? Is the trend in most plots real global warming or is there a modeling issue in ERA5? Please estimate the amplitude of the monument thermoelastic signal at these sites for comparison.

Response 7:

Thank you for this helpful comment. We agree that the original figure did not provide enough context for interpreting the selected stations. In the revised manuscript, we have revised the figure caption and related text to specify that the examples are daily FEM_{FL} TED time series at four representative GNSS stations.

We also agree that apparent trend-like variations in these model series should not be interpreted as evidence of global warming. In the revised manuscript, we no longer discuss TED trends as a main result. We clarify that low-frequency variations in the modelled TED series may reflect variability or uncertainty in the ERA5 soil-temperature forcing and other model inputs, and that linear trends fitted to the TED series should be interpreted with caution.

Regarding monument thermoelastic effects, we have added an order-of-magnitude comparison between subsurface TED and above-ground monument thermal expansion in the revised manuscript. This comparison is now included in Fig. 5 and discussed in Section 3.1. Because complete information on monument height, material, geometry, anchoring condition, and local thermal environment is only available for a subset of GNSS stations, the site-level estimates are treated as an order-of-magnitude comparison rather than as part of the released TED correction product.

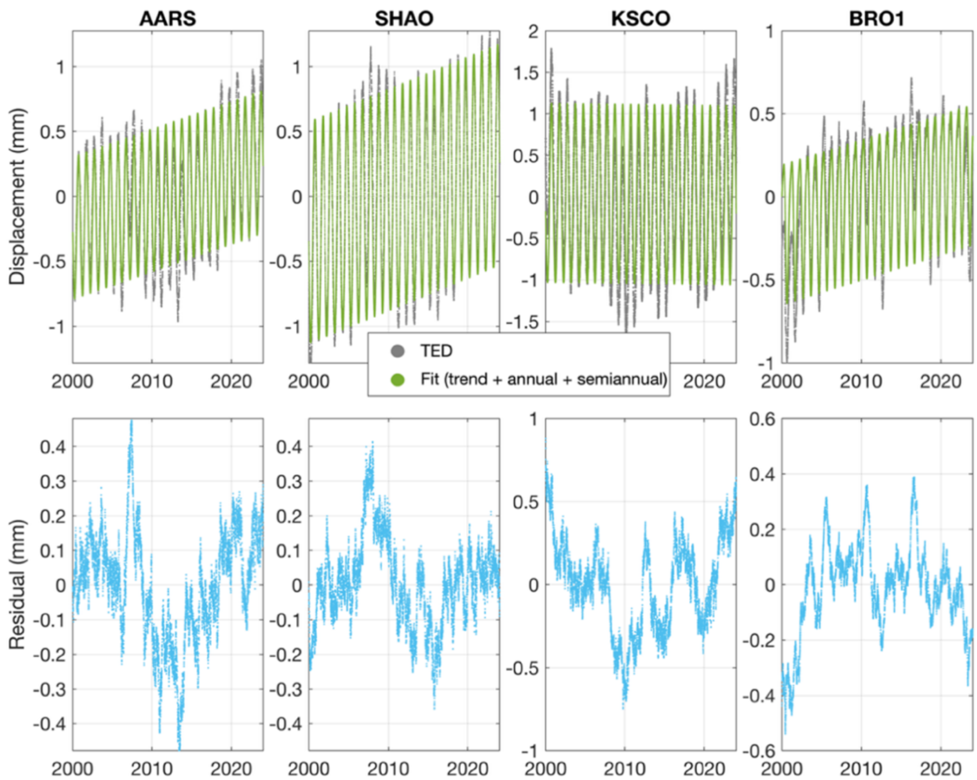

Figure 4 Examples of vertical TED time series at four representative GNSS stations. The upper panels show the daily FEM_{FL} TED series together with the fitted linear, annual, and semiannual components. The lower panels show the residuals after removing these fitted components.

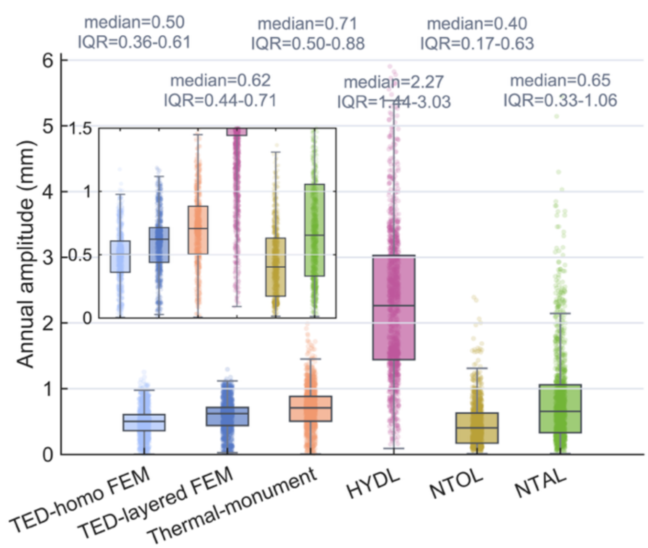

Figure 5 Order-of-magnitude comparison of annual vertical displacement amplitudes from TED, monument thermal expansion, and non-tidal loading at 2,612 GNSS stations with documented monument information. TED-FEM denotes the traditional homogeneous FEM solution, while TED-FEM_{FL} denotes the FEM_{FL} solution used for the

released TED product. Thermal-monument represents above-ground monument thermal expansion estimated from the available monument information. HYDL, NTOL, and NTAL denote hydrological, non-tidal oceanic, and non-tidal atmospheric loading displacements, respectively.

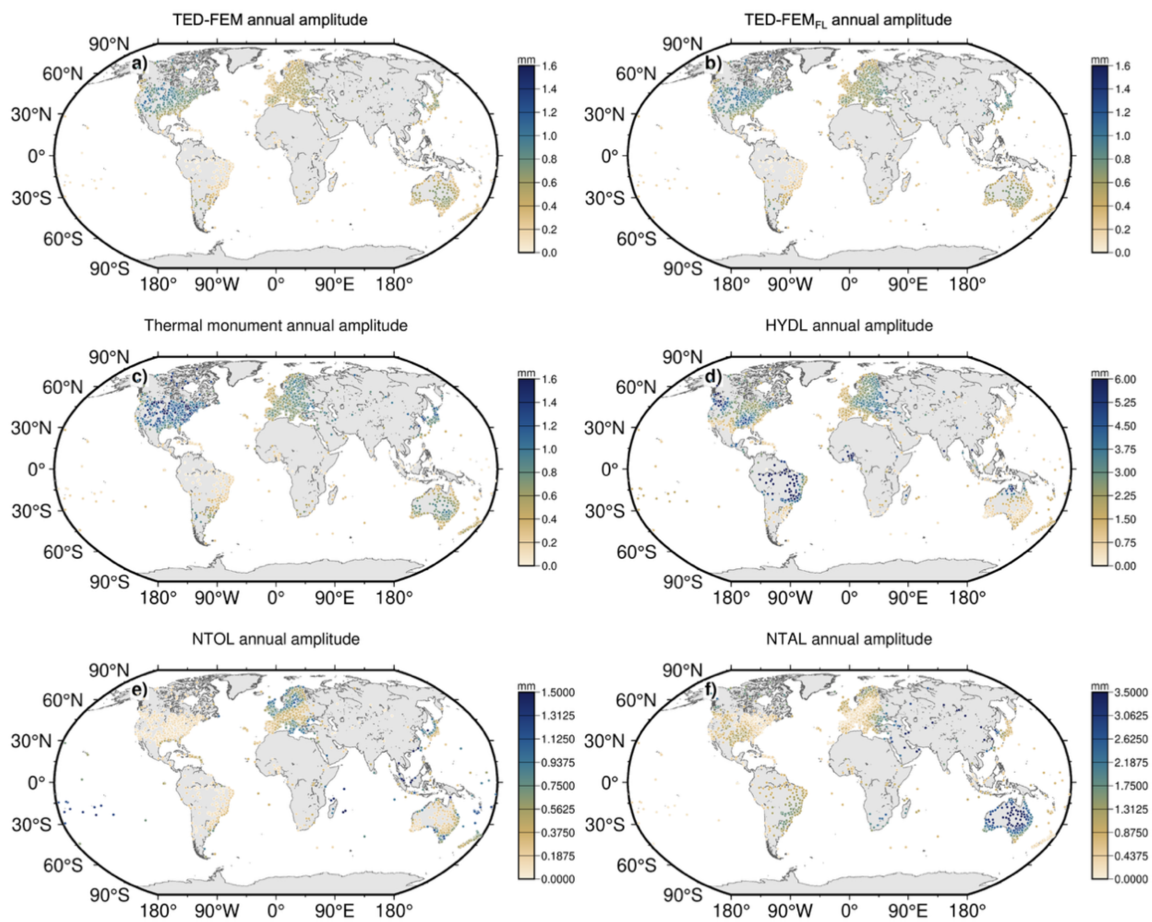

Figure. Spatial comparison of annual vertical displacement amplitudes from TED, monument thermal expansion, and non-tidal loading. Panels (a) and (b) show the annual amplitudes of the traditional homogeneous FEM solution and the FEMFL TED product, respectively. Panel (c) shows above-ground monument thermal expansion, and panels (d)–(f) show hydrological loading (HYDL), non-tidal ocean loading (NTOL), and non-tidal atmospheric loading (NTAL), respectively.

Comment 8: Figure 5 peak to peak differences looks very similar to Figure 4 annual differences. Please comment.

Response 8: Thank you for this comment. We agree that the spatial pattern of the percentile-based peak-to-peak TED variability is similar to that of the annual TED amplitude. This similarity is expected because the annual component is the dominant contribution to TED signal at most stations. However, the two quantities describe different aspects of the model output. The annual amplitude is obtained from harmonic fitting and represents only the seasonal annual component, whereas the percentile-based peak-to-peak metric is calculated from the full daily TED series as the difference between the 95th and 5th percentiles. It therefore includes the annual, semiannual, and nonseasonal components while reducing the influence of occasional extreme values.

We have added a clarification in the revised manuscript to explain this relationship. The similarity between the two figures indicates that the annual cycle controls much of the total TED variability, while local differences between them reflect additional non-annual variability.

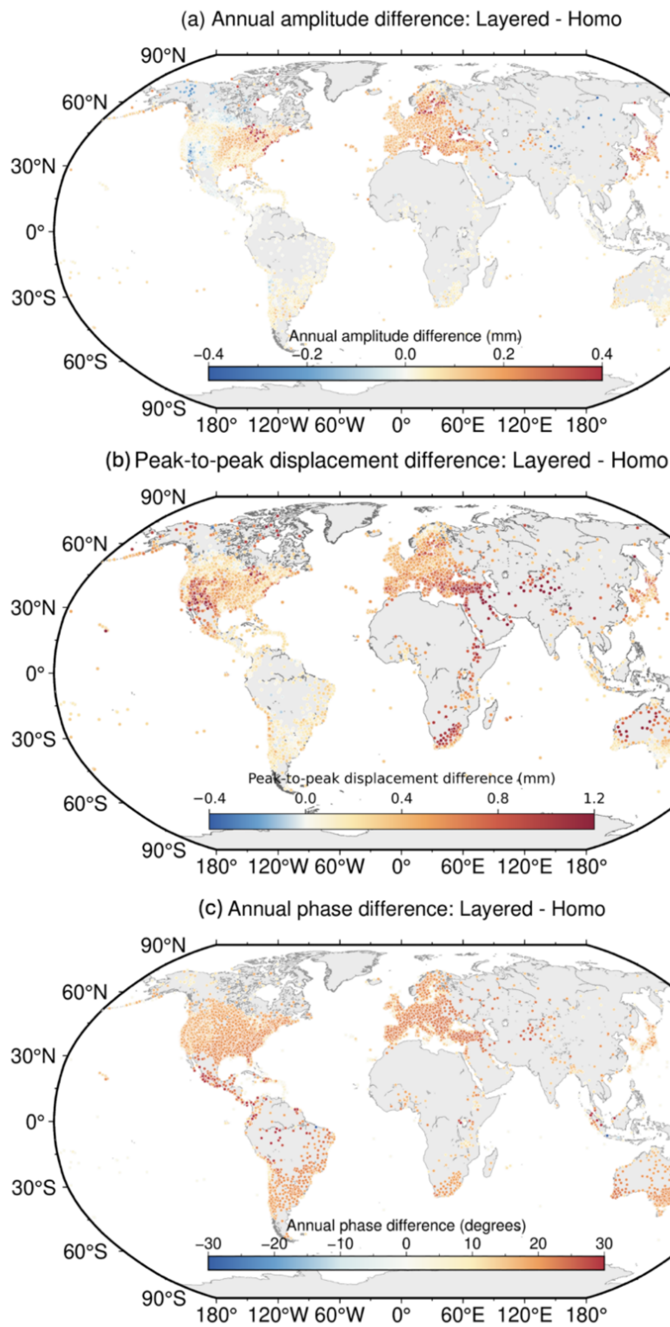

Figure 6 Station-based differences between the FEM_{FL} TED product and the traditional homogeneous FEM solution. Panels (a)–(c) show the differences in annual amplitude, percentile-based peak-to-peak displacement, and annual phase, respectively.

Comment 9: Linear trends in the TED model. A reader cannot interpret figure 7 because we cannot see the lengths of the vectors when there are many sites. Again, it would useful to calculate TED over an equally spaced grid on the continents and use a color legend.

1. ERA5 is expected to contain trends due to global warming. The trends seem to range from +- 2 deg. C per decade (see Evaluating global and regional land warming trends in the past decades with both MODIS and ERA5-Land land surface temperature data, Wang et al, 2022). A discussion of the trends in the soil temperatures in the ERA5 data should be discussed.
2. The discussion should include a discussion of the uncertainty in the soil temperature trends.

Response 9:

We agree that the original trend-vector figure was difficult to interpret because the dense station distribution made the vector lengths hard to read. We also agree that interpreting linear trends in the TED model would require a separate analysis of ERA5 soil-temperature trends and their uncertainties.

In the revised manuscript, we have removed the TED linear-trend figure and the related trend-focused discussion. The revised paper now focuses on the annual, semiannual, and nonseasonal TED variability, which are the main intended applications of the released correction product. We do not interpret fitted linear trends in the TED series as evidence of long-term geophysical deformation or global warming.

Besides, to avoid ambiguity, we have added a cautionary statement in the Data availability section that linear trends fitted to the TED series should be interpreted with caution because they may be affected by uncertainties in the long-term behaviour of the temperature forcing and model inputs. The linear term in the harmonic model is retained only as part of the fitting procedure for estimating seasonal and nonseasonal components, and it is not presented as a reliable long-term geophysical trend.

Comment 10: Figure 8 does not add anything to the paper. The authors are comparing the half-space model and the layered model in a slightly different way. However, at this point the reader gets the point that the models are different and produce different results. What we want to know is the validity of any of the models.

Response 10:

We agree that the previous figure mainly repeated the comparison between the homogeneous FEM solution and the FEM_{FL} solution, without directly addressing their performance against GNSS observations. We have therefore removed this figure from the main text and shortened the related discussion.

In the revised manuscript, we refocused this part on GNSS-based WRMS assessment. Section 3.3 now first presents the TED-only WRMS changes, and we also compare the WRMS changes obtained from the FEM_{FL} correction and the traditional homogeneous FEM correction. This provides an observation-based assessment of the practical correction performance of the two TED models, rather than another model-to-model comparison.

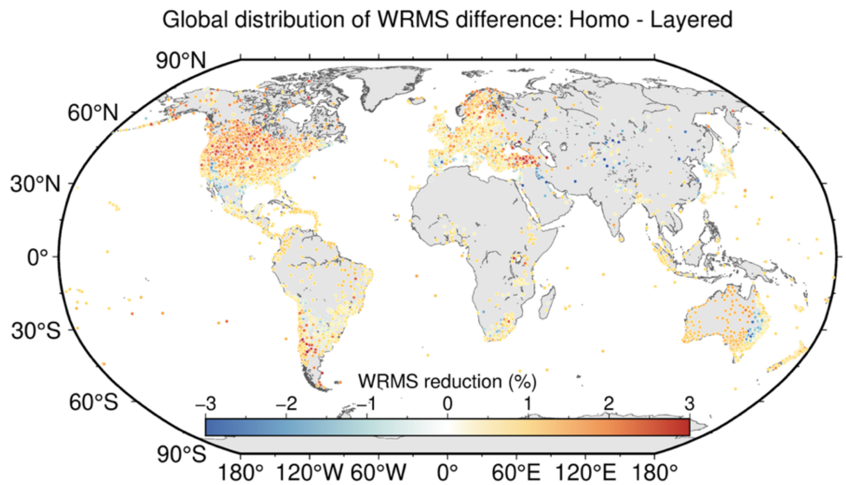

Figure 12 Difference in WRMS reduction between the FEM_{FL} TED correction and the traditional homogeneous FEM correction. Positive values indicate a larger WRMS reduction from the FEM_{FL} correction.

Comment 11: Figure 9 and its discussion is about validating the SoilGrids in the FEM_{FL} It is not about evaluating the robustness of the FEM_{FL} model. In both cases you are using the FEM_{FL} model and comparing output using 1) SoilGrids as input and 2) site geology information as input.

Response 11:

We agree that this analysis does not directly evaluate the robustness of the FEMFL model itself, because both calculations use the same FEMFL framework. Rather, it evaluates the influence of the subsurface input data by comparing FEMFL solutions generated with two different parameter sources: SoilGrids-based parameters and available site-information-constrained parameters.

In the revised manuscript, we have changed the wording accordingly. We no longer describe this analysis as a validation or robustness test of the FEM_{FL} model. Instead, it is now presented as an assessment of the SoilGrids-based subsurface parameterization and its influence on the modelled TED estimates. The figure caption and the related discussion in Section 3.2 have also been revised to make this distinction clear.

Revised text:

"Figure 7 examines the effect of using SoilGrids as global source of shallow subsurface information. For the 2,612 stations with documented foundation-depth information, FEM_{FL} solutions generated with SoilGrids-based parameters are compared with solutions constrained by the available site information. The resulting differences are small: annual-amplitude differences are mostly about 0.01–0.03 mm, and annual-phase differences are generally within about 1–3°. Overall, these differences are much smaller than the typical annual TED signal, suggesting that SoilGrids-based parameterization introduces limited uncertainty in the modelled annual cycle for these stations."

Comment 12: No comments on figures 10-11. I feel like these figures do not help the reader gain insight into the physics of the problem.

Response 12: In the revised manuscript, we removed the previous figures and discussion that repeated the homogeneous–layered model comparison without adding substantial new insight. The Results section has been shortened and refocused on analyses that are more directly relevant to the released dataset, including the order-of-magnitude comparison, the sensitivity to subsurface parameterization and model parameters, spectral diagnostics, and the TED-only effect on GNSS height time series.

Comment 13: The discussion starting at L556 regarding the WRMS reduction is important and should be highlighted in the abstract. L570-571 in particular are very important and should be stated in the abstract. This is what the readers want to know.

Response 13:

Thank you for this important comment. We agree that the standalone WRMS reduction from the TED correction is one of the key results that readers need to know early in the manuscript. In the revised manuscript, we have therefore highlighted the TED-only WRMS result in the Abstract.

Following the reviewer's earlier comments, we have removed the previous statement based on the combined TED + non-tidal loading correction. The Abstract now reports only the effect of applying TED alone to GNSS height time series. We also revised Section 3.3 to present this TED-only WRMS assessment at the beginning of the GNSS-based assessment, rather than near the end of the manuscript.

Revised text in the Abstract:

"When applied alone to GNSS height time series, TED produces positive WRMS reductions at about 79% of the assessed stations, with most improvements within 3–5% and the largest WRMS reductions reaching about 10%."

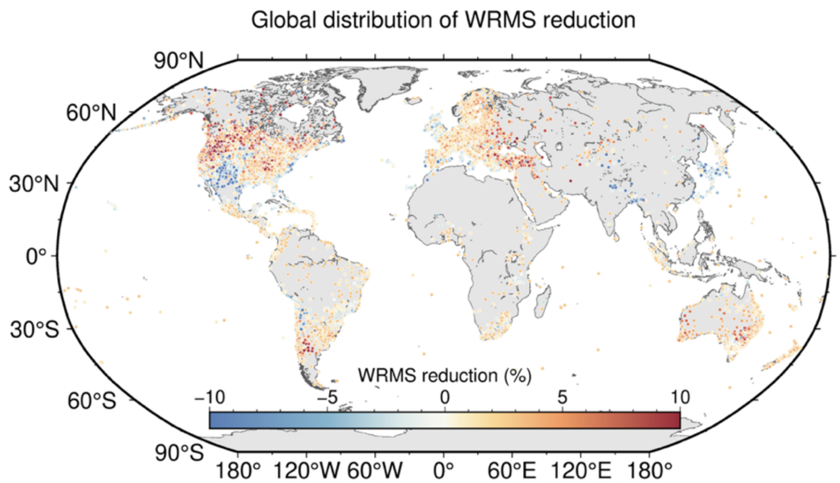

Figure 11 Global distribution of WRMS changes after applying the TED-only correction to GNSS height time series. Positive values indicate reduced WRMS after removing TED, while negative values indicate increased WRMS.