The manuscript presents a potentially valuable global dataset of monthly soil and vegetation component temperatures derived from the FuSVeT framework, with broad spatial coverage and clear relevance for land–atmosphere studies. The study is promising and may make a useful contribution to the remote sensing community. At the same time, several important aspects would benefit from further clarification and strengthening, particularly with respect to methodological novelty, the physical interpretation of the retrieved variables, model applicability across different climatic regimes, parameterization assumptions, validation design, the independence assumption in triple collocation, and uncertainty analysis. Addressing these points would help improve the overall rigor of the study and enhance confidence in the conclusions. For these reasons, I believe the manuscript would benefit from major revision before further consideration:

Reply:

We sincerely appreciate the time and effort you devoted to reviewing our manuscript, as well as your positive evaluation of the potential value of GloSVeT for the remote sensing and land–atmosphere research communities. Your comments and suggestions are highly constructive and have helped us improve the methodological clarity, physical interpretation, validation design, and overall rigor of the manuscript.

We have carefully considered each of your comments and made substantial revisions accordingly. The major revisions are summarized as follows:

1. We clarified the methodological novelty of the present dataset paper and the physical meaning/applicability boundaries of the retrieved component temperatures (Q1 and Q2).
2. We supplemented the discussion of the applicability of the GOT09-based MDC model in polar regions and clarified the rationale for using Bayesian optimization (Q3 and Q4).
3. We added an emissivity sensitivity analysis to assess the influence of the fixed emissivity assumption and included an internal closure check to examine the consistency of the decomposition–reconstruction process (Q5 and Q8).
4. We revised potentially misleading statements related to the in situ validation strategy and trend significance interpretation (Q6 and Q9).
5. We updated the dataset combination used in the TC analysis to reduce potential error dependence and added further justification for the use of physically related proxy variables in the TC framework (Q7).

Detailed point-by-point responses are provided below. For ease of review, all revisions in the revised manuscript are highlighted in red.

Q1. The main retrieval method (FuSVeT) is adopted from a previously published study (Liu et al., 2025, RSE). In its current form, the manuscript appears to primarily extend the application of this existing approach to a longer temporal span, rather than introducing substantial methodological advancements. The authors should justify the novelty and added value of the present study for the remote sensing community.

Reply:

Thank you for this valuable comment. We agree that the core FuSVeT retrieval framework was developed in our previous study (Liu et al., 2025). The objective of the present manuscript is therefore not to claim a new retrieval algorithm, but to extend FuSVeT from method development to the generation, validation, and dissemination of a long-term global component temperature dataset.

As a data descriptor, the novelty and added value of this study lie mainly in three aspects. First, to our knowledge, GloSVeT is the first global dataset that simultaneously provides monthly mean surface soil and vegetation component temperatures at 0.05 °spatial resolution over a multi-decadal period from 2003 to 2023. This dataset provides a new observational basis for separating soil and vegetation thermal dynamics in land–atmosphere interaction and ecosystem studies. Second, this study implements FuSVeT at the global scale through multi-source data integration and large-scale automated processing, which is a necessary step for transforming the previously developed retrieval framework into a usable community dataset. Third, we established a comprehensive evaluation strategy tailored to component temperature products, including an internal closure check, site-based validation, TC analysis, indicator-based physical consistency assessment, and examination of seasonal spatiotemporal consistency. This evaluation framework helps address the long-standing difficulty of validating soil and vegetation component temperatures in the absence of direct global ground truth.

To clarify the novelty and added value of the present study, we revised the Introduction as follows: “To address these challenges, this study presents GloSVeT, the first global dataset of monthly mean surface soil and vegetation component temperatures from 2003 to 2023 at 0.05 °

spatial resolution. The dataset is generated using the previously developed FuSVeT framework, which is expanded from algorithm development to global implementation through multi-source integration and large-scale automated processing. Beyond dataset production, this study further establishes a comprehensive evaluation strategy tailored to component temperature products, combining internal closure check, site-based validation against in situ temperature observations from a comprehensive network, TC analysis incorporating reanalysis soil and air temperatures, indicator-based consistency assessment with soil moisture and solar-induced fluorescence (SIF), and examination of seasonal spatiotemporal consistency. This multi-perspective assessment helps address the long standing difficulty of evaluating component temperatures in the absence of direct global ground truth.” (Lines 104-112)

Q2. What is the physical definition of the monthly mean soil and vegetation temperatures? The MxD11C3 LST product is derived under clear-sky conditions, whereas ERA5 LST incorporates the effects of cloud cover. It is therefore unclear whether the retrieved component temperatures implicitly include cloud-related information. How are the effects of sunlit and shaded soil surfaces addressed? A simple averaging of clear-sky LST may inevitably mix sunlit and shaded signals, potentially introducing bias.

Reply:

Thank you for this professional comment. As also discussed in our response to Reviewer #1 Q3, the target variable of GloSVeT is not a clear-sky monthly mean component temperature, but an estimate of monthly mean component temperatures that is as close as possible to the all-weather monthly mean thermal state. Physically, the retrieved monthly mean component temperatures represent effective radiometric temperatures of the visible soil and vegetation components within the sensor field of view, averaged over the monthly mean diurnal cycle.

We agree that MODIS LST observations are fundamentally constrained by clear-sky conditions, whereas ERA5-Land includes all-sky land surface thermal information. In the GloSVeT framework, ERA5-Land hourly data are used to provide continuous all-sky diurnal-cycle information and to help constrain the monthly mean diurnal cycle parameters. Therefore, ERA5-Land helps reduce the temporal sampling limitation and clear-sky sampling effect associated with MODIS observations. However, the retrieved component temperatures are still ultimately constrained by clear-sky TIR observations, and thus the mismatch between clear-sky MODIS constraints and the intended all-weather monthly mean target remains an important source of uncertainty. Regarding sunlit and shaded surfaces, the current global 0.05° product adopts a two-component framework that separates the mixed-pixel LST into effective soil and vegetation component temperatures. It does not explicitly distinguish sunlit and shaded fractions within either component. Therefore, sunlit–shaded thermal contrasts, especially for the soil component, may introduce additional uncertainty. We have explicitly added this limitation in the revised manuscript.

The relevant revisions are provided below.

Influence of clear-sky sampling (Lines 668-678)

A second limitation arises from the clear-sky nature of TIR satellite observations and uncertainties in related input data. The MODIS LST observations used in GloSVeT are available only under clear-sky conditions, whereas the intended monthly mean component temperatures are expected to represent the average thermal state under all-weather conditions within each month. Although ERA5-Land provides continuous all-sky diurnal thermal information and helps constrain the monthly mean diurnal cycle, the retrieved component temperatures are still ultimately constrained by clear-sky TIR observations. This mismatch between clear-sky observation and the all-weather monthly mean target may introduce sampling biases, especially in persistently cloudy regions or during periods when cloudy-sky and clear-sky surface thermal conditions differ systematically. Such uncertainty could be reduced by incorporating cloud-contaminated LST reconstruction algorithms (Jia et al., 2024; Du et al., 2025) or by directly using advanced all-weather LST products. For example, GHA-LST (Jia et al., 2023), a global, hourly, 5 km, and all-weather LST dataset, provides a promising data source for improving the temporal completeness and all-weather representativeness of component temperature retrievals.

Limitation related to sunlit and shaded fractions (Lines 654-667)

Second, the current two component framework approximates the land surface as a combination of vegetation and soil components … This simplification also does not explicitly distinguish between sunlit and shaded fractions, although their temperature differences can be substantial, especially for the soil component (Ermida et al., 2014) … Future developments could further integrate more detailed surface component structure information ...

Q3. The GOT09 model is employed as the MDC model in this study. However, this model may not be applicable in polar regions, where polar day and night occur. It is unclear how the FuSVeT method is applied in regions above 66.5 °N, as shown in Fig. 3. Do these regions exhibit the same monthly diurnal cycle characteristics as lower latitudes? Further clarification is needed.

Reply:

Thank you for this valuable comment. We agree that the application of diurnal temperature cycle models in polar regions is limited, as also discussed by Hong et al. (2021). The same limitation applies to the GOT09-based MDC model used in FuSVeT/GloSVeT, because this model assumes a relatively regular diurnal temperature cycle characterized by daytime warming and nighttime cooling. This assumption may be weakened or violated during polar day and polar night, when the conventional diurnal heating and cooling cycle becomes weak or absent.

However, this does not mean that FuSVeT/GloSVeT assumes the same monthly diurnal-cycle characteristics in polar regions as in lower-latitude regions. To further examine this issue, we conducted a parameter diagnostic analysis using the fitted diurnal amplitude parameter T_a in June and December 2023. As shown in Figure R2, the latitudinal distributions of T_a exhibit distinct high-latitude behavior near and above the Arctic Circle. In particular, the fitted amplitudes and their interquartile ranges change markedly in polar-day and polar-night affected regions, indicating that the fitted MDC parameters can reflect high-latitude differences rather than simply following the broad mid-latitude pattern.

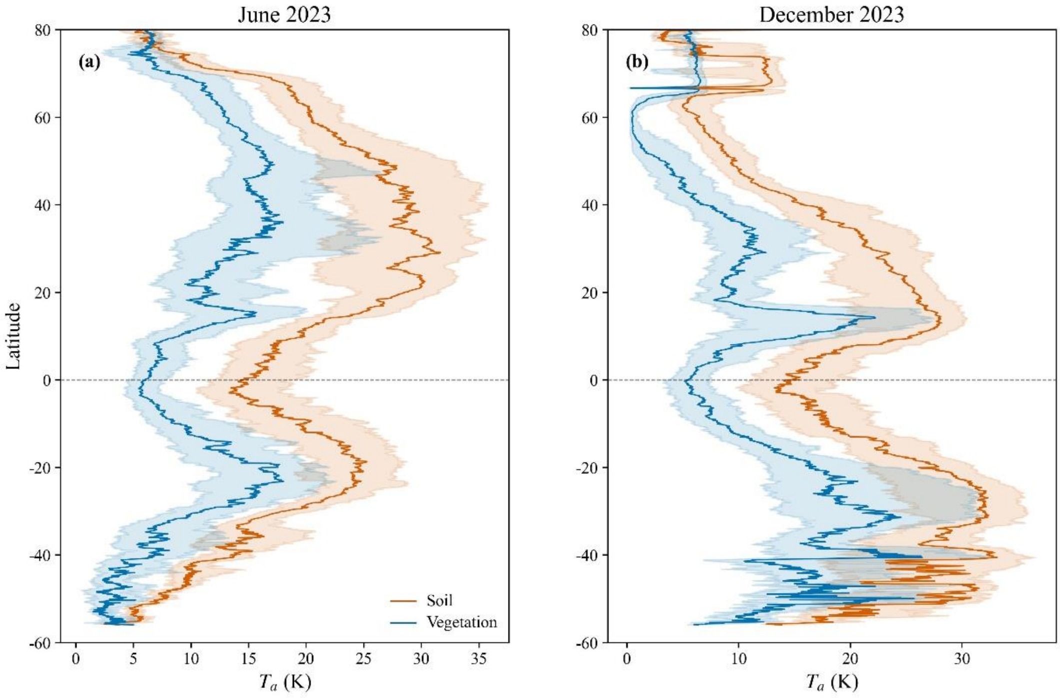

Figure R2. Latitudinal distributions of the fitted diurnal amplitude parameter T_a for soil and vegetation temperatures in June and December 2023. Solid lines indicate zonal median values, and shaded areas indicate the 25th–75th percentile ranges.

Nevertheless, we agree that the physical interpretation and reliability of DTC-based retrievals become more uncertain under polar-day and polar-night conditions. Therefore, in the revised manuscript, we explicitly clarified this limitation and reminded users that GloSVeT retrievals in high-latitude regions should be interpreted with caution. The revised text reads: "Finally, GloSVeT relies on the GOT09-based MDC model, which assumes a relatively regular diurnal temperature cycle characterized by daytime warming and nighttime cooling. This assumption can be weakened or violated in high-latitude regions during polar day and polar night, when the conventional diurnal heating and cooling cycle becomes weak or absent (Hong et al., 2021). Under such conditions, GloSVeT retrievals may have larger uncertainties and should be interpreted with caution. Developing diurnal temperature cycle formulations that are better suited to polar illumination conditions and weak diurnal temperature cycles would help enhance the applicability of GloSVeT in these climatically sensitive regions." (Lines 696-701).

Q4. The manuscript states that the number of iterations for Bayesian optimization is empirically set to 2000, but this choice lacks sufficient justification. Does the optimal number of iterations vary with land cover type? In addition, it would be helpful to quantify the computational cost and explain why this approach is preferred over simpler alternatives such as widely used Levenberg–Marquardt minimizationin DTC modeling. The description of "approximating the solution distribution" should also be clarified for better precision.

Reply:

We appreciate this professional and helpful comment. We have revised the corresponding description and added further justification for the choice of Bayesian optimization and the iteration number. The reason for using Bayesian optimization rather than the conventional Levenberg–Marquardt minimization has been discussed in Liu et al. (2020). As shown in the screenshot below, which is taken from Figure 8 of Liu et al. (2020), Bayesian optimization achieved higher retrieval accuracy than the Levenberg–Marquardt method for the soil and vegetation component temperature separation problem. This is mainly because the optimization in FuSVeT/GloSVeT involves nonlinear, constrained, and underdetermined parameter estimation. In particular, the soil and vegetation temperatures are not directly optimized parameters, but are derived from the fitted model parameters; therefore, imposing a physically reasonable soil–vegetation temperature difference constraint is not straightforward in a conventional Levenberg–Marquardt framework. Bayesian optimization provides a more flexible way to search the constrained parameter space and is generally more robust for this type of underdetermined nonlinear problem.

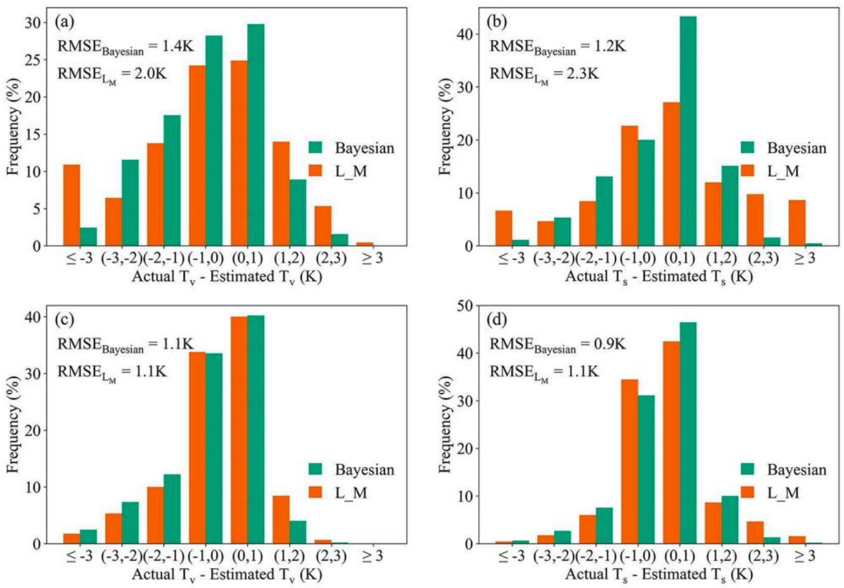

Fig. 8. Histograms of the differences between the estimated and actual component temperatures. (a) Vegetation temperature for FVC STD = 0.05, (b) soil temperature for FVC STD = 0.05, (c) vegetation temperature for FVC STD = 0.1, and (d) soil temperature for FVC STD = 0.1.

To justify the iteration number, we conducted a convergence test across different FVC bins, because FVC directly controls the relative contributions of soil and vegetation components in the mixed-pixel thermal signal and is therefore more closely related to the difficulty of component temperature separation than categorical land cover type. Specifically, using the July 2023 data, valid pixels were stratified into ten FVC bins, and 1000 samples were randomly selected from each bin. Bayesian optimization was then performed with 500, 1000, 1500, 2000, 3000, 5000, 7000, and 9000 iterations, and the 10000-iteration result was used as the reference solution. The MAE of the retrieved soil and vegetation component temperatures relative to this reference was then calculated. As shown in Figure A1, after 2000 iterations, further increasing the iteration number led only to gradual improvements. Since the computational cost of Bayesian optimization increases approximately with the number of iterations, using 2000 iterations requires only about 20% of the optimization cost of the 10000-iteration reference while retaining stable retrieval performance. Therefore, we selected 2000 iterations as a practical balance between retrieval stability and computational efficiency.

We also revised the imprecise expression “approximating the solution distribution”. The revised description now states that Bayesian optimization needs to iteratively build a probabilistic surrogate of the objective function and sequentially sample promising regions of the parameter space to identify the parameter set that minimizes the reconstruction error.

The revisions are listed below.

Why Bayesian optimization was used instead of the Levenberg–Marquardt method (Lines 254-257)

To ensure physical realism, a temperature difference constraint between soil and vegetation components (−10 K to 30 K) was imposed. Considering this constraint and the need for robust parameter estimation, the parameters were solved using a Bayesian optimization algorithm rather than the conventional Levenberg–Marquardt method (Liu et al., 2020).

Why 2000 iterations were used (Lines 275-282)

Second, Bayesian optimization is computationally intensive since it needs to iteratively build a probabilistic surrogate of the objective function and sequentially sample promising regions of the parameter space to identify the parameter set that minimizes the reconstruction error (Pelikan, 2005). To determine an appropriate iteration number, we conducted a convergence test using the July 2023 data across different FVC bins. The result from 10000 iterations was used as the reference solution, and the retrievals obtained with fewer iterations were compared against this reference. As shown in Figure A1, after 2000 iterations, further increasing the iteration number led only to gradual improvements. Therefore, the iteration number was set to 2000 as a practical balance between retrieval stability and computational efficiency.

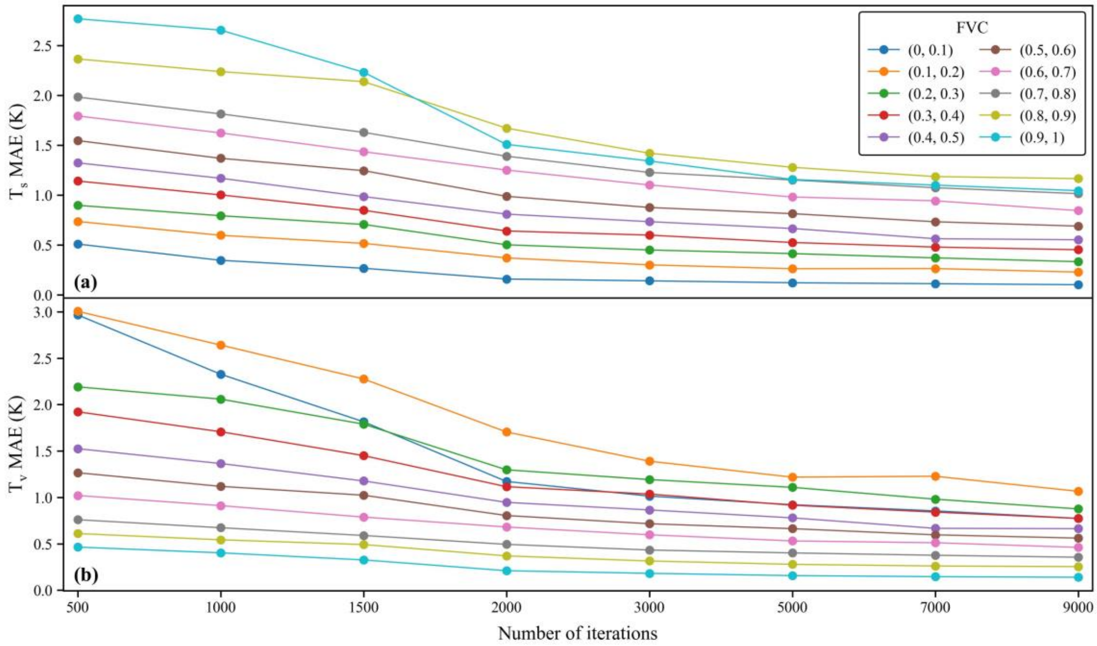

Figure A1. Convergence test of Bayesian optimization across FVC bins. Using the July 2023 data, valid pixels were stratified into ten FVC bins, and 1000 samples were randomly selected from each bin. Bayesian optimization was performed with 500, 1000, 1500, 2000, 3000, 5000, 7000, and 9000 iterations, and the result from 10000 iterations was used as the reference solution. The MAE values of retrieved soil temperature (a) and vegetation temperature (b) were calculated relative to this reference solution. Different lines represent different FVC bins. (Lines 744-748)

Q5. The emissivity in the MxD11C3 product is retrieved using a day/night algorithm, whereas in the FuSVeT method, soil and vegetation emissivities are fixed globally at 0.95 and 0.98 for the entire period (2003–2023). This assumption may not be sufficiently realistic. For example, in bare soil regions (e.g., the Sahara), where vegetation cover is negligible, the emissivity should be consistent with the MxD11C3 product rather than a fixed value. This issue should be revisited and better justified.

Reply:

We sincerely appreciate your valuable comment. We agree that the use of fixed broadband emissivities for soil and vegetation is a simplifying assumption and may introduce uncertainty for a global product, especially over bare soils, deserts, drylands, and heterogeneous mineral backgrounds.

In the early stage of this study, we examined the broadband emissivity characteristics of globally distributed pure pixels by using the emissivity information from MxD11C3 together with narrowband-to-broadband conversion relationships. The results indicated that the broadband emissivities of pure soil and vegetation pixels were generally close to 0.95 and 0.98, respectively. Therefore, these two values were adopted as the baseline emissivity settings in GloSVeT. It should also be noted that the emissivity provided by MxD11C3 represents a pixel-level effective emissivity retrieved from the day/night algorithm, whereas GloSVeT requires component-level broadband emissivities for soil and vegetation. Therefore, directly assigning the MxD11C3 pixel-level emissivity to soil and vegetation components is not straightforward, particularly for mixed pixels.

Nevertheless, we agree that the globally fixed emissivity assumption remains an important source of uncertainty. To address this concern, we added a new discussion subsection entitled "Influence of fixed emissivity" and conducted an emissivity sensitivity analysis. Specifically, nine soil–vegetation emissivity combinations were tested, with soil emissivity set to 0.90, 0.93, and 0.95, and vegetation emissivity set to 0.97, 0.98, and 0.99. Based on the 2023 monthly datasets, global valid pixels were stratified into FVC bins at an interval of 0.1, and 1000 valid samples were randomly selected from each FVC bin for each month. The component temperatures retrieved under the other eight emissivity perturbation scenarios were then compared with the baseline setting used in GloSVeT (ε_s = 0.95 and ε_v = 0.98), and the resulting MAE was quantified.

The results indicate that the fixed-emissivity assumption has a clear FVC-dependent influence. Without separating FVC conditions, the median MAEs of soil and vegetation component temperatures were 1.16 K and 1.26 K, respectively, suggesting that the overall influence remains within a relatively limited range. More importantly, the sensitivity analysis reveals that emissivity uncertainty has a smaller impact on the dominant component within a mixed pixel. When FVC is below 0.5, the soil component dominates the mixed-pixel thermal signal, and the MAE of soil temperature is generally below 1 K. Conversely, when FVC exceeds 0.5, the vegetation component is more strongly constrained, and the MAE of vegetation temperature is generally below 1 K. Larger uncertainties mainly occur for the weakly represented component under very low or very high FVC conditions, where the MAE can approach or exceed 2 K. We therefore clarified that the fixed-emissivity assumption does not substantially weaken the overall ability of GloSVeT to represent the dominant component temperature, but the non-dominant component temperature under extreme FVC conditions should be interpreted with caution.

For clarity, the revision is listed below. (Lines 603-629)

5.1 Influence of fixed emissivity

In GloSVeT, the broadband emissivities of soil and vegetation were uniformly set to 0.95 and 0.98, respectively. For a global product, such fixed emissivity assumptions inevitably introduce uncertainty, particularly for soil component (García-Santos et al., 2015). Therefore, we conducted an emissivity sensitivity analysis to evaluate the potential influence of this assumption on the retrieved soil and vegetation component temperatures. Specifically, nine soil–vegetation emissivity combinations were tested, with soil emissivity set to 0.90, 0.93, and 0.95, and vegetation emissivity set to 0.97, 0.98, and 0.99. Based on the 12 monthly datasets in 2023, the global valid pixels were stratified into FVC bins at an interval of 0.1, and 1000 valid samples were randomly selected from each FVC bin for each month. Soil and vegetation component temperatures were then recalculated under each emissivity combination. Taking the baseline setting used in GloSVeT (ε_s = 0.95 and ε_v = 0.98) as the reference, the mean absolute error (MAE) induced by the other eight emissivity perturbation scenarios was calculated.

Figure 13 presents the distribution of MAE across FVC bins, where each boxplot contains 96 MAE statistics derived from 12 months and 8 emissivity perturbation scenarios. The results reveal that sensitivity to emissivity assumptions is strongly dependent on FVC: as FVC increases, the emissivity sensitivity of soil component temperature gradually increases, whereas that of vegetation component temperature decreases. Without separating FVC conditions, the median MAEs of T_s and T_v are 1.16 K and 1.26 K, respectively, suggesting that the overall influence of the fixed emissivity assumption remains within a relatively limited range. A more detailed FVC-dependent analysis further indicates that emissivity uncertainty has a smaller impact on the dominant component within a mixed pixel. When FVC is below 0.5, the soil component dominates the mixed-pixel thermal signal, and the MAE of T_s is generally below 1 K. Conversely, when FVC exceeds 0.5, the vegetation component is more strongly constrained, and the MAE of T_v is typically below 1 K. In contrast, under very low or very high FVC conditions, the weakly represented component contributes less to the mixed-pixel thermal signal and therefore becomes more sensitive to emissivity perturbations. Therefore, the fixed emissivity assumption does not substantially weaken the overall ability of GloSVeT to represent the dominant component temperature, but the non-dominant component temperature under extreme FVC conditions should be interpreted with caution.

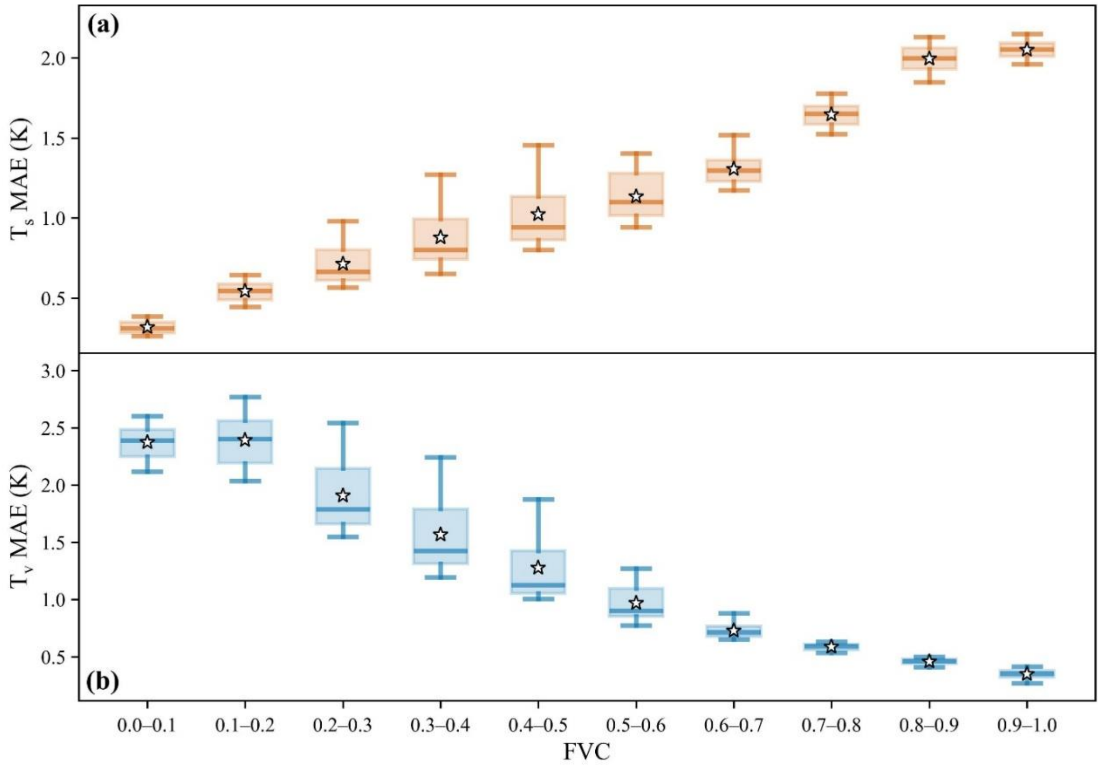

Figure 13: Sensitivity of GloSVeT to emissivity perturbations across FVC gradients for soil temperature (a) and vegetation temperature (b). Boxplots summarize the MAE distributions derived from 12 monthly datasets in 2023 and eight emissivity perturbation scenarios relative to the baseline emissivity setting used in GloSVeT.

Q6. For in situ validation, sites classified as purely soil- or vegetation-covered based on NDVI are used as references. However, this approach raises concerns. The LST of a pure pixel essentially corresponds to the component temperature itself, meaning that the validation may effectively assess monthly mean LST rather than true component temperatures. In addition, the number of validation samples appears limited (e.g., only 66 points for soil temperature validation in summer during 2003–2023). The authors should justify the dataset used and consider incorporating dedicated LSCT validation sites (e.g., EVO, KAL, and DM).

Reply:

We appreciate this valuable comment. We apologize that our original description was not sufficiently clear. We agree that tower-derived radiometric LST cannot be directly interpreted as component temperature under mixed surface conditions. Therefore, in this study, it was used as an approximate component temperature reference only under near-pure soil or near-pure vegetation conditions. Under these conditions, the footprint-scale LST is dominated by one component and can reasonably serve as a reference for the corresponding component temperature within the 0.05° GloSVeT pixel. The limited number of validation records results from the combined use of spatial representativeness screening, and NDVI-based near-pure surface filtering. Although this strict screening reduces the number of retained records, it improves their suitability for component temperature validation. In addition, the 66 summer soil temperature records are monthly matched records over 2003–2023, corresponding to several years of monthly observations rather than a small number of instantaneous samples. Therefore, the number of validation samples is not as limited as it appears. We also agree that dedicated LSCT sites such as EVO, KAL, and DM would be highly valuable because they can provide simultaneous soil, vegetation, and even sunlit/shaded component temperatures. However, these datasets are not currently openly available for our long-term global validation, so they were not included in this study.

In the revised manuscript, we have clarified the validation logic as follows: "This tower-derived radiometric LST was used as an approximate component temperature reference only when the tower footprint was dominated by either bare soil or vegetation. Typically, the effective observational footprints of these flux towers range from 40 to 80 m in diameter. To identify near pure surface conditions, we additionally used the Landsat Collection 2 Tier 1 Level-2 8-day 30 m normalized difference vegetation index (NDVI) composite (https://developers.google.com/earth-engine/datasets/catalog/LANDSAT_COMPOSITES_C02_T1_L2_8DAY_NDVI). Specifically, the 8-day NDVI values at each site were aggregated to monthly means. The corresponding tower-derived LST records were used as soil temperature references when NDVI ≤ 0.1 and as vegetation temperature references when NDVI ≥ 0.7, whereas records with intermediate NDVI values were excluded from the component temperature validation." (Lines 181-188).

Q7. The manuscript employs Triple Collocation (TC) to evaluate consistency among datasets; however, this method relies on the assumption of independent errors. In this study, ERA5 data are used both as input for retrieving soil and vegetation temperatures and as a reference dataset (e.g., ERA5 soil temperature), which may violate the independence assumption and introduce correlated errors. Please clarify how this issue is addressed. Furthermore, the use of subsurface ERA5-Land soil temperature (0–7 cm) may not be consistent with the penetration depth of thermal infrared signals. As noted by the authors, air temperature is also not an ideal proxy for vegetation temperature. These choices require further justification.

Reply:

This comment is highly appreciated. We agree that the independence assumption is critical for the TC analysis. Although ERA5-Land skin temperature and ERA5-Land soil temperature are different variables, they are generated within the same land surface modelling and forcing framework. Therefore, their errors may not be fully independent, which could weaken the independence assumption required by TC. To address this issue, we updated the TC triplet in the revised manuscript by excluding ERA5-Land from the TC analysis and using JRA-3Q instead. JRA-3Q was selected because it provides globally continuous land surface analysis variables over the study period and is produced by an independent reanalysis system that was not used in the generation of GloSVeT. In addition, because GLDAS was already included in the TC triplet, JRA-3Q was preferred over other potential alternatives such as MERRA-2, since MERRA-2 and GLDAS are both NASA products. Although this does not completely guarantee error independence, it helps reduce potential dependence among the datasets in terms of data source, modelling system, and institutional origin.

We also agree that the comparison variables used in TC are not identical to the GloSVeT component temperatures. The reanalysis soil temperatures represent near-surface soil thermal states at specific depths, whereas GloSVeT retrieves a TIR-constrained surface radiometric temperature. Similarly, near-surface air temperature is not equivalent to vegetation temperature, even though it is physically coupled with canopy temperature in densely vegetated regions. Therefore, in the revised manuscript, we have clarified that JRA-3Q and GLDAS were used as physically related proxy datasets for TC-based relative consistency evaluation, rather than as ground-truth references. Correspondingly, we revised Section 4.2 “TC-based uncertainty evaluation”, added a justification for the selected reanalysis variables in the Data section, and discussed the remaining limitations of proxy-based comparison.

For clarity, the major revisions are listed below.

Reason for selecting the reanalysis datasets (Lines 189-204):

Because long-term, observation-based soil and vegetation temperature datasets with global coverage are currently unavailable, two reanalysis datasets, JRA-3Q (Kosaka et al., 2024) and GLDAS (Rodell et al., 2004), were used as physically related comparison datasets for the TC analysis. The JRA-3Q product provides monthly averaged variables at a spatial resolution of 0.375 °, while GLDAS offers 0.25 °resolution. To maximize consistency with the GloSVeT soil temperature, only the top-layer soil temperature was used from both datasets. Specifically, the monthly mean soil temperature at level 1 layer (0-2 cm) from JRA-3Q and the SoilTMP0_10cm_inst variable (0-10cm) from GLDAS were employed. Although these variables differ in their effective soil depth and are not identical to the TIR surface soil component temperature retrieved by GloSVeT, they are physically related near-surface soil thermal variables and can therefore serve as reasonable proxy datasets for evaluating relative consistency within the TC framework. For vegetation temperature, near-surface air temperature was used as a proxy in densely vegetated regions, given the strong coupling between canopy temperature and near-surface air temperature through land–atmosphere heat exchange (Li et al., 2001; Zhan et al., 2011; Rutter et al., 2023). Accordingly, the monthly mean 2 m air temperature from JRA-3Q and the Tair_f_inst variable from GLDAS were used for the vegetation-temperature TC analysis. For consistency in spatial support, GloSVeT soil and vegetation temperatures, GLDAS soil temperature and air temperature, and GLASS FVC were aggregated to the 0.375 °JRA-3Q grid using an area-weighted method. The aggregated GLASS FVC was further used to identify densely vegetated pixels ($FVC \geq 0.8$) for the evaluation of vegetation temperature.

Revised TC-based uncertainty evaluation (Lines 383-463):

4.3 TC-based uncertainty evaluation

Figure 4 shows the results of the TC-based correlation (R) for soil temperature. In general, the three products exhibit largely consistent spatial patterns (Figures 4a–c) and latitudinal variation (Figure 4d). High correlations ($R > 0.7$) are found mainly in the Northern Hemisphere high latitudes and the humid tropics, whereas low correlations ($R < 0.5$) appear in many mid-latitude transitional regions. This spatial coherence across independent datasets indicates that the major large-scale thermal variations are consistently captured among the different data sources, and further supports the physical plausibility of the GloSVeT retrievals.

Although their patterns are broadly similar, some differences can be observed among the three datasets. Overall, the inter-product differences are more evident in the Northern Hemisphere than in the Southern Hemisphere (Figure 4d). JRA-3Q shows the highest correlations over regions north of 30 °N, but shows the weakest performance across much of the remaining areas, particularly in tropical regions between 10 °N and 30 °N (Figure 4d). Its distribution contains a larger share of both very high ($R > 0.8$) and very low ($R < 0.4$) correlations (Figures 4e, f), reflecting spatial heterogeneity and less stable behaviour at global scale. In contrast, GLDAS and GloSVeT achieve more equivalent performance (Figures 4e, f). GLDAS shows a distribution concentrated around moderate-to-high R values (R between 0.6 and 0.8), whereas GloSVeT presents a smoother distribution and maintains more concentrated distribution toward higher R values (Figure 4a, e). In particular, GloSVeT shows locally higher or comparable R values in humid tropical and subtropical regions, such as southern China, the Amazon and the Congo Basin, underscoring its advantage in these environmentally sensitive and data-scarce areas. By contrast, its correlations are lower at high latitudes. This may occur because JRA-3Q and GLDAS represent layer-averaged subsurface soil temperatures, whereas GloSVeT soil temperature represents the radiometric temperature of the visible surface and is more sensitive to snow cover and radiative inversions (Cao et al., 2023; Liu et al., 2025). Taken together, the TC-based evaluation suggests that reanalyses perform better in high latitude regions, whereas GloSVeT shows a relatively more balanced performance across latitudes and retains certain advantages in humid tropical regions. These results indicate that GloSVeT provides reliable surface soil temperature estimates and serves as a useful complement to conventional reanalysis soil temperature products for large-scale applications.

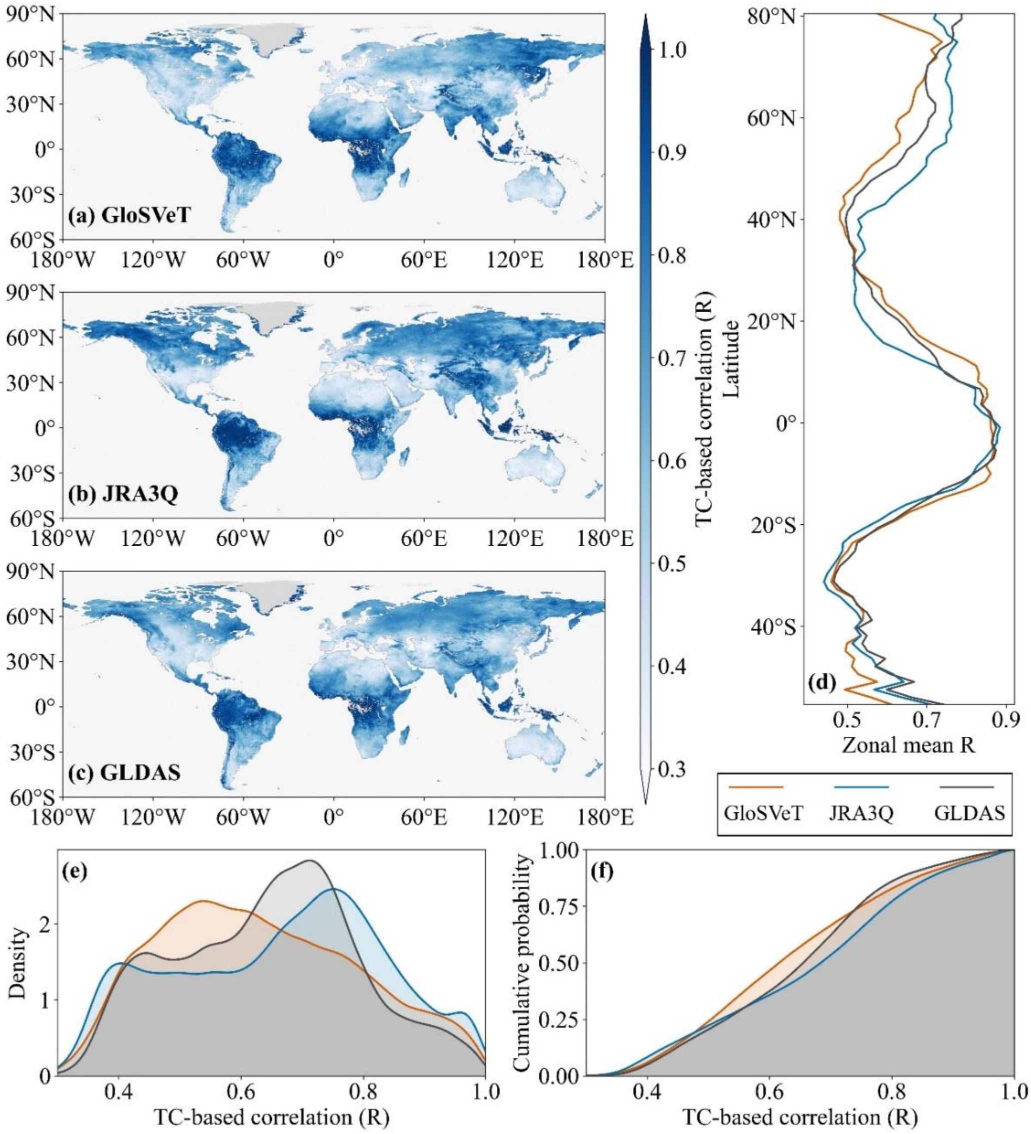

Figure 4: TC-based correlation (R) for soil temperature. Panels (a–c) illustrate the spatial distributions for GloSVeT, JRA-3Q, and GLDAS, respectively, with light gray and medium gray indicating the ocean and land backgrounds; (d) shows the zonal-mean R as a function of latitude; and panels (e) and (f) present the probability density and cumulative distribution functions, respectively.

To extend the assessment from a global view to ecosystem contexts, this study further compared TC-based correlations for soil temperature across 14 biome types (Figure 5). In general, the three datasets show higher median correlations (R ≈ 0.7–0.9) in cold high-latitude and humid tropical ecosystems, such as temperate coniferous and mixed forests, boreal forests, tundra (Figure 5e, 5f, 5k), and tropical moist broadleaf forests and mangroves (Figure 5a, 5n). In contrast, lower correlations (R ≈ 0.5) with larger spreads are observed in Mediterranean, arid, and temperate grassland systems (Figure 5l, 5m, 5h). This cross-biome consistency reinforces that large-scale soil thermal variations are physically coherent across independent datasets.

Among the three products, JRA-3Q shows the largest inter-biome variability. It performs best in cold and mid-latitude forest ecosystems, such as boreal forests, tundra, montane grasslands, and temperate mixed or coniferous forests (Figure 5f, 5k, 5j, 5d, 5e), but its correlations drop markedly in tropical and semi-arid regions. Its boxplots also display wider interquartile ranges, indicating higher spatial heterogeneity and weaker internal consistency. By contrast, GloSVeT and GLDAS exhibit more comparable performance across most biomes. GLDAS generally shows moderate-to-high correlations, whereas GloSVeT maintains relatively balanced correlations across biome types and performs particularly well in tropical dry broadleaf forests, tropical/subtropical grasslands, savannas and shrublands, flooded grasslands, deserts/xeric shrublands, and Mangrove (Figure 5b, 5g, 5i, 5m, 5n), highlighting its ability to capture surface soil thermal variability under complex vegetation and moisture conditions. Overall, the biome-specific comparison supports the global results in Figure 4, indicating that reanalysis products tend to perform better in cold and high-latitude ecosystems, whereas GloSVeT provides a relatively robust performance across diverse biome types with advantages in humid tropical and moisture-variable environments.

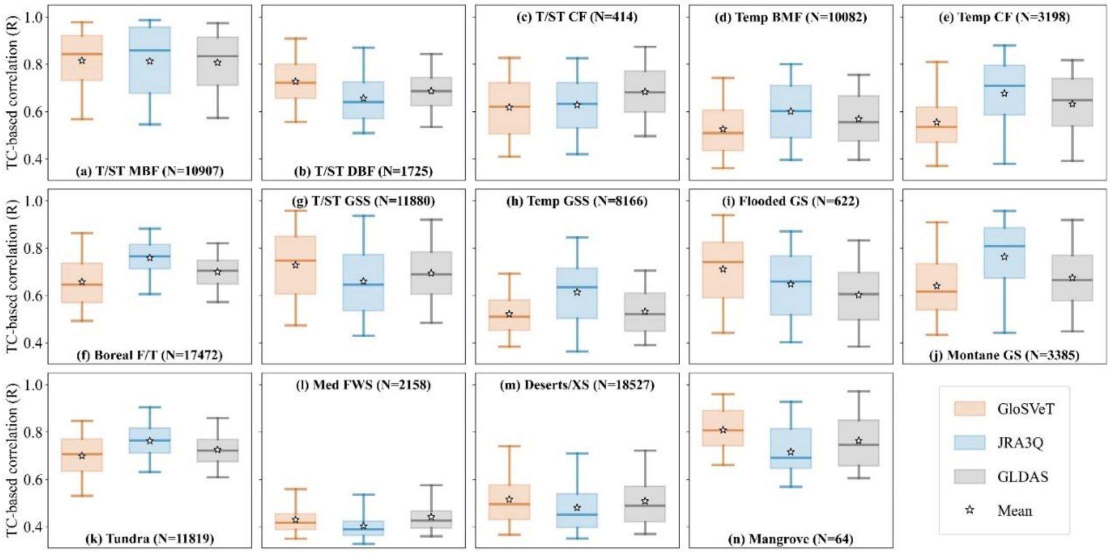

Figure 5: TC-based correlation (R) for soil temperature across 14 major biomes. Boxplots show the distributions of R for GloSVeT (orange), JRA-3Q (blue), and GLDAS (dark gray). Stars denote the mean values. T/ST, MBF, DBF, CF, Temp, BMF, F/T, GSS, GS, Med, FWS, and XS, are the abbreviations of 'Tropical/subtropical', 'moist broadleaf forests', 'dry broadleaf forests', 'coniferous forests', 'Temperate', 'broadleaf and mixed forests', 'forests/taiga', 'grasslands, savannas and shrublands', 'grasslands and savannas', 'Mediterranean', 'forests, woodlands and scrub or sclerophyll forests', and 'xeric shrublands', respectively.

Figure 6 and Figure 7 present the TC-based correlations for vegetation temperature. Since the analysis was restricted to densely vegetated pixels, and only pixels with more than 100 matched pairs were included to ensure the reliability of TC estimation, the resulting coverage was concentrated mainly over tropical and subtropical regions. Similar to the soil component, the three datasets display broadly consistent spatial and latitudinal patterns, yet GloSVeT systematically exhibits higher correlations and a narrower spread. As shown in Figures 6a–d, across South America, Africa, and Southeast Asia, GloSVeT achieves stronger and more spatially uniform correlations ($R > 0.8$) than JRA-3Q and GLDAS. The probability density and cumulative distribution curves (Figures 6e,f) further indicate that GloSVeT contains a greater fraction of high R pixels and fewer low R outliers, reflecting improved temporal coherence and spatial stability. When compared across major vegetation types (Figure 7), GloSVeT shows the highest or comparable median correlations in tropical/subtropical moist broadleaf forests, tropical/subtropical dry broadleaf forests, and temperate broadleaf and mixed forests (Figure 7a–c). In tropical/subtropical grasslands, JRA-3Q exhibits slightly higher median correlations, but GloSVeT remains close to JRA-3Q and clearly higher than GLDAS (Figure 7d). These results underscore the superior performance of GloSVeT in representing canopy thermal dynamics across diverse hydroclimatic conditions.

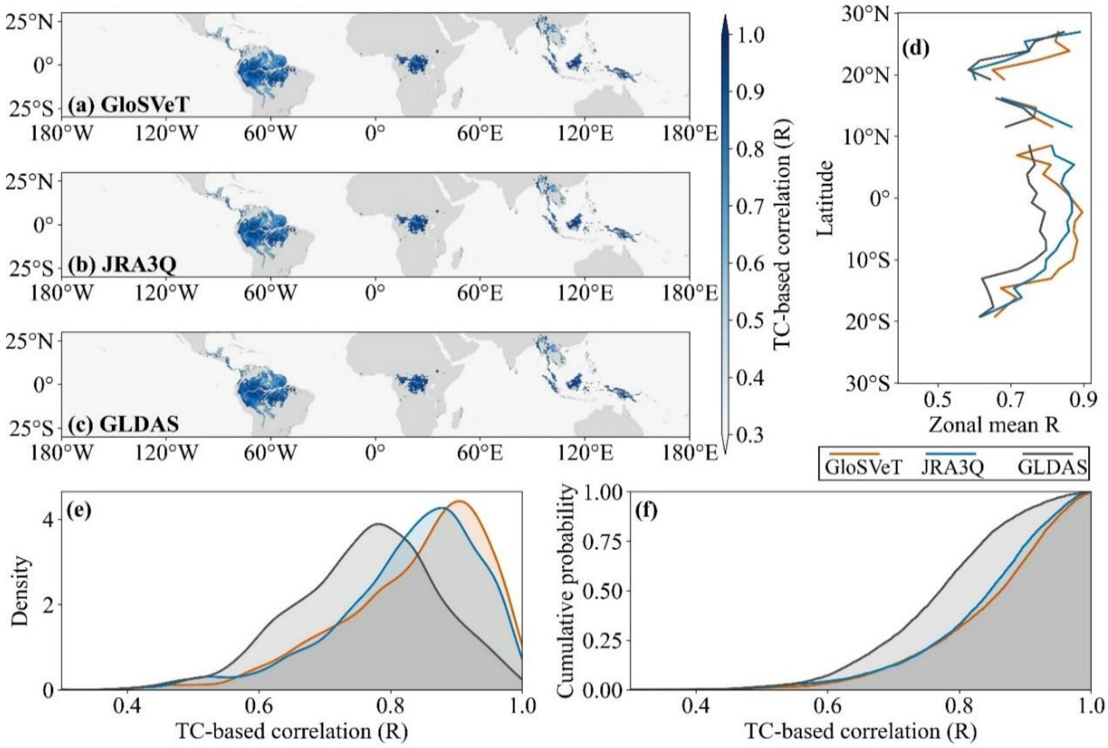

Figure 6: TC-based correlation (R) for vegetation temperature. Panels (a–c) illustrate the spatial distributions for GloSVeT, JRA-3Q, and GLDAS, respectively, with light gray and medium gray indicating the ocean and land backgrounds; (d) shows the zonal-mean R as a function of latitude; and panels (e) and (f) present the probability density and cumulative distribution functions, respectively.

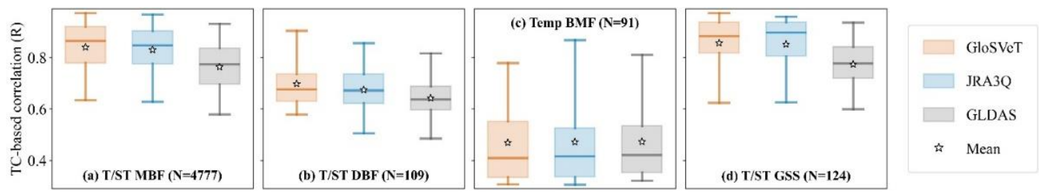

Figure 7: TC-based correlation (R) for vegetation temperature across four major biomes. Boxplots show the distributions of R for GloSVeT (orange), JRA-3Q (blue), and GLDAS (dark gray). Stars denote the mean values. T/ST, MBF, DBF, Temp, BMF, and GSS, are the abbreviations of 'Tropical/subtropical', 'moist broadleaf forests', 'dry broadleaf forests', 'Temperate', 'broadleaf and mixed forests', and 'grasslands, savannas and shrublands', respectively.

Discussion of limitations (Lines 685-695):

The validation strategy also has inherent limitations. Although the TC analysis provides an effective way to evaluate GloSVeT in the absence of globally available long-term observations of soil and vegetation temperatures, its results should be interpreted within the context of proxy-based comparison. The reanalysis soil temperatures used in TC represent near-surface soil thermal states at specific depths, which are physically related to but not identical to the TIR-constrained surface radiometric temperature retrieved by GloSVeT (Cao et al., 2023; Liu et al., 2025). Similarly, near-surface air temperature was used as an auxiliary proxy for vegetation temperature in densely vegetated regions (Li et al., 2001; Zhan et al., 2011; Rutter et al., 2023), but it is not equivalent to vegetation temperature. Therefore, the TC-based correlations mainly indicate the relative temporal consistency between GloSVeT and physically related independent datasets, rather than absolute accuracy against true soil or vegetation component temperatures. More direct and spatially representative observations of surface soil and vegetation temperatures would help further strengthen the independent validation of GloSVeT and reduce the uncertainty associated with proxy-based comparisons.

Q8. Reanalysis data are extensively used in the FuSVeT framework, along with MxD11C3 LST. The fusion of multiple datasets may lead to error accumulation; however, uncertainty propagation is not analyzed in the manuscript. Additionally, the reliability of using the MDC model independently should be assessed, as errors introduced during separate fitting may affect the results. For example, the reconstructed 24-hour monthly mean LST could be validated against geostationary LST data at the pixel level.

Reply:

Thank you for this comment. We agree that the use of multiple data sources in FuSVeT/GloSVeT may introduce uncertainties from different parts of the workflow. The purpose of introducing reanalysis data, however, is to provide continuous diurnal cycle information and reduce the dimensionality of MDC parameter estimation, thereby alleviating the underdetermination caused by the limited number of MODIS LST retrievals. In this study, instead of conducting a full step by step uncertainty propagation analysis, we evaluated the final target variables of GloSVeT, namely monthly mean soil and vegetation component temperatures, through multiple complementary approaches, including in situ validation, TC analysis, physical consistency assessment, and emissivity sensitivity analysis.

Following your suggestion and that of Reviewer #1 Q8, we also added an internal closure check to examine whether the retrieved component temperatures can preserve the original mixed pixel LST signal. Specifically, the retrieved soil and vegetation component temperatures were recombined into synthetic mixed pixel LSTs at the four MODIS overpass times and compared with the corresponding original MODIS LST retrievals. The results show that the monthly global RMSEs generally ranged from approximately 1.2 to 2.0 K, and the bias values mostly remained within about ±1 K during 2003–2023. This indicates that the decomposition and reconstruction process is internally consistent and does not introduce large structural deviations from the original MODIS thermal constraints.

Regarding the independent assessment of the MDC model, we agree that a dedicated comparison using geostationary LST data would be valuable. However, a full independent evaluation of different DTC/MDC formulations is beyond the scope of this dataset paper. Previous studies (Duan et al., 2012; Hong et al., 2018) have already conducted detailed comparisons among different DTC models, and the GOT09 model was adopted in this study based on its robust performance in these prior evaluations.

The added description of the internal closure check is listed below. **(Lines 325-350)**

4.1 Internal closure check

Figure 2 summarizes the internal closure check for GloSVeT. Overall, the reconstructed mixed-pixel LSTs were reasonably consistent with the original MODIS LST values at the four overpass times. The monthly global RMSEs generally ranged from approximately 1.2 to 2.0 K, while the bias values mostly remained within about ±1 K. The relatively compact boxplots further suggest that the closure errors were temporally stable during 2003–2023, with RMSE variations generally within about 0.2–0.5 K and bias variations within about 0.4–0.6 K.

The closure residuals also exhibited clear overpass-time dependence. In terms of RMSE, MYD11C3_Nighttime (i.e., 01:30 nominal local time) yielded the smallest closure error, with a median RMSE of about 1.24 K, whereas MOD11C3_Daytime (i.e., 10:30 nominal local time) had the largest error, with a median RMSE close to 1.94 K. MYD11C3_Daytime (i.e., 13:30 nominal local time) and MOD11C3_Nighttime (i.e., 22:30 nominal local time) had intermediate RMSEs, both around 1.61 K. The bias statistics revealed a broadly consistent overpass-dependent pattern. MOD11C3_Daytime and MYD11C3_Nighttime had negative median biases of -1.04 and -0.64 K, respectively, whereas MYD11C3_Daytime and MOD11C3_Nighttime had positive median biases of 0.92 and 0.56 K, respectively. Although the biases were not uniformly negative, the negative residuals were slightly larger in magnitude than the positive residuals, suggesting a weak overall negative tendency in the closure residuals. This overpass-dependent behavior may be partly related to the uneven temporal distribution of MODIS sampling times (Liu et al., 2025). In particular, the long interval between the MYD11C3_Nighttime and MOD11C3_Daytime overpasses spans the nocturnal minimum, sunrise transition, and morning warming period, but this key transition phase is not directly constrained by MODIS. As a result, the fitted MDC curve may still contain residual uncertainties in phase and morning warming dynamics, which can lead to overpass-dependent closure biases. Overall, the closure check suggests that GloSVeT can largely reproduce the original MODIS mixed-pixel LST signal, while systematic overpass-dependent residuals remain.

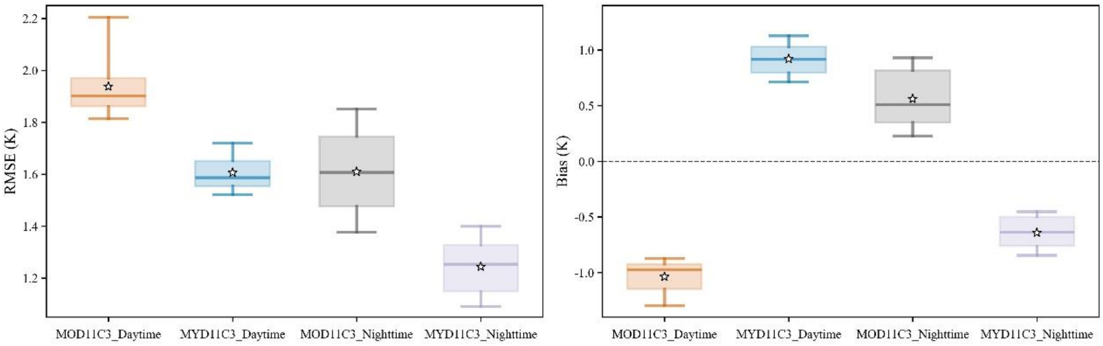

Figure 2: Internal closure check of GloSVeT. Boxplots show the distributions of monthly global (a) RMSE and (b) Bias between reconstructed mixed-pixel LST and the corresponding original MODIS LST during 2003–2023 at the four MODIS overpass times. Stars denote the mean values, and the dashed line in panel (b) indicates zero Bias.

References

[1] Duan, S.-B., Li, Z.-L., Wang, N., Wu, H., and Tang, B.-H.: Evaluation of six land-surface diurnal temperature cycle models using clear-sky in situ and satellite data, Remote Sens. Environ., 124, 15-25, 10.1016/j.rse.2012.04.016, 2012.

[2] Hong, F., Zhan, W., Göttsche, F.-M., Liu, Z., Zhou, J., Huang, F., Lai, J., and Li, M.: Comprehensive assessment of four-parameter diurnal land surface temperature cycle models under clear-sky, ISPRS J. Photogramm. Remote Sens., 142, 190-204, 10.1016/j.isprsjprs.2018.06.008, 2018.

Q9. Figures 10–11 indicate that only a subset of pixels passes the statistical significance test ($p < 0.05$). The authors should report the proportion and spatial distribution of statistically significant pixels and discuss the robustness of their conclusions, given that a considerable fraction of the results may not be significant. Furthermore, the potential impact of non-significant regions on the overall interpretation should be addressed.

Reply:

Thank you for this helpful comment. We have revised the discussion of Figures 10–11 (Figures 11-12 in the new manuscript) by explicitly reporting the proportions of statistically significant pixels and clarifying their spatial distribution. We also added a statement that non-significant regions should be interpreted as areas with weak or uncertain pixel-level trends, rather than as evidence of no thermal change. Accordingly, we softened the conclusion from "confirm" to "suggest" and clarified that the figures support large-scale spatial tendencies and statistically robust trend centres, while pixel-level interpretations in non-significant regions should be made with caution.

For your convenience, the revisions are listed below. **(Lines 567-591)**

To further investigate the spatial heterogeneity of these temporal trends, the linear slopes were mapped for representative cold (January) and warm (July) months (Figures 11–12). In January, the majority of pixels showed positive trends, accounting for 68.4% of soil temperature pixels and 62.3% of vegetation temperature pixels, indicating a widespread tendency toward wintertime warming. Pronounced warming occurs across North America, Europe, and southeastern China, whereas localized cooling is evident in northern Eurasia and the high Asian interior (Bartusek et al., 2022; Qiao et al., 2025). However, only 6.7% of soil temperature pixels and 7.0% of vegetation temperature pixels passed the 0.05 significance test. These significant pixels were not randomly distributed, but were mainly concentrated in regions with relatively strong warming or cooling signals, indicating that the most pronounced trend centres are more statistically robust. The large fraction of non-significant pixels is likely related to the limited number of annual samples for each calendar month (21 values during 2003–2023) and strong interannual variability. Therefore, the non-significant regions should be interpreted as areas with weak or uncertain pixel-level trends, rather than as evidence of no thermal change. Compared with soil temperature, vegetation trends are generally weaker, consistent with the effects of snow masking, stronger canopy–air coupling during cold seasons, and the thermal buffering imposed by canopy structure and aerodynamic resistance (Forzieri et al., 2017). In the Southern Hemisphere, January trends are weaker and more heterogeneous, reflecting the smaller land area, weaker continental thermal contrast, and stronger oceanic modulation during austral summer. In July, warming becomes more spatially extensive covering 74.8% of soil temperature and 70.7% of vegetation temperature pixels, but the overall rate is slightly lower. Broad warming is observed across the Northern Hemisphere mid-latitudes, particularly in western North America, the Mediterranean–Middle East, Central Asia, and northeastern China, whereas parts of Europe display moderate cooling. The proportion of statistically significant pixels increased in July, reaching 16.6% for soil temperature and 10.8% for vegetation temperature, indicating a stronger and more spatially coherent warm-season trend signal than in January. Consistent with the cold-season behaviour, soil temperature trends exceed those of vegetation temperature, reflecting the greater sensitivity of the soil skin to net radiation and the reduced evaporative buffering from vegetation (Qiao et al., 2023). In summary, these distinct seasonal behaviours, including stronger winter contrasts at high latitudes and widespread summer warming over continental interiors, suggest that GloSVeT captures meaningful large-scale structures of surface thermal change, while pixel-level interpretations in non-significant regions should be made with caution.