Review of:

GloSVeT: a global 0.05 °monthly mean surface soil and vegetation component temperature dataset (2003-2023)

This review is structured: summary of the study and conclusion, major considerations, minor considerations.

Summary:

This study presents a new dual-dataset product called GloSVeT, which provides both soil and vegetation temperatures. Results are validated against 8 in situ sites, as well as ERA5-Land and GLDAS using triple collocation. There is also separate analysis of the results by biome, including trend analysis. The author's conclusions are that this dataset demonstrates high correlations, high accuracy, and shows evidence of physical consistency.

I would recommend major considerations before further consideration.

Reply:

Thank you for your careful review of our manuscript and for your constructive comments and suggestions. We have carefully considered all of your comments and revised the manuscript accordingly. The main revisions are summarized as follows:

1. We updated the dataset combination used in the TC analysis and revised the corresponding analyses and interpretations (Q1).
2. We revised inappropriate statements or added further explanations to avoid ambiguity and potential misunderstanding (Q2-Q4, Q8, Q10, Q12-Q16, Q22-Q25, and Q27).
3. We improved the presentation and wording of figures and figure captions to make them clearer and more self-contained (Q17, Q18, Q20, and Q26).
4. We removed unnecessary content that was not used in this study (Q5).
5. We supplemented methodological and descriptive details where additional clarification was needed (Q6, Q7, Q9, Q11, Q19, and Q21).

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

Major Considerations:

Q1. ERA5-Land Skin Temperature is used for the generation of GloSVeT, and ERA5-Land Top-layer soil temperature and ERA5-Land Air temperature are used for the evaluation. Are these datasets suitably independent for triple collocation? I would like to read a justification of this.

Reply:

We greatly appreciate this important comment. We agree that the independence assumption is critical for the TC analysis. Although ERA5-Land skin temperature, top-layer soil temperature, and 2 m air 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 revised the TC triplet by excluding ERA5-Land from the TC analysis and using JRA-3Q instead. The revised TC analysis was therefore conducted using GloSVeT, JRA-3Q, and GLDAS. 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 choice does not completely guarantee error independence, it helps reduce potential dependence among the datasets in terms of data source, modelling system, and institutional origin.

Correspondingly, we revised Section "TC-based uncertainty evaluation" and updated the related figures and interpretations based on the new triplet. The revised text has been highlighted in red in the manuscript (please see Lines 383-463). For convenience, the main revision is also listed below.

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.

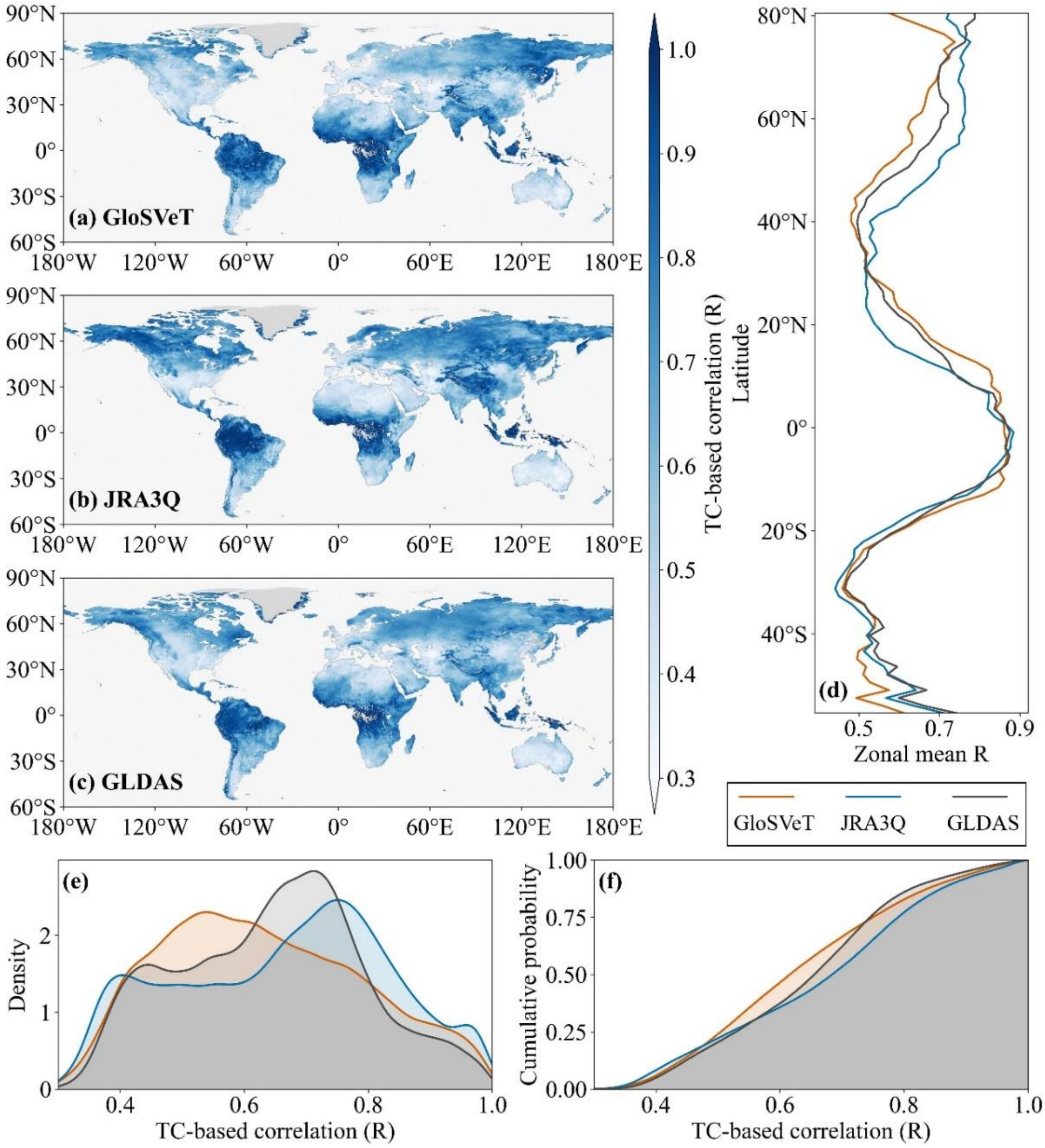

Figure 4: TC-based correlation (R) for soil temperature. Panels (a–c) illustrate the spatial distributions for GloSVeT, JRA-3Q, and GLDAS, respectively, with light gray and medium gray indicating the ocean and land backgrounds; (d) shows the zonal-mean R as a function of latitude; and panels (e) and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

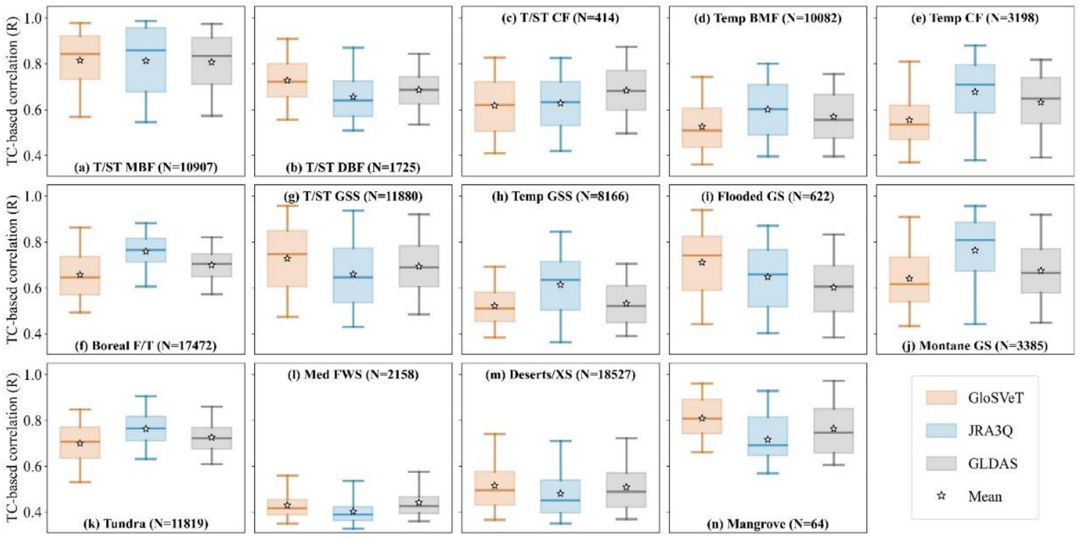

Figure 5: TC-based correlation (R) for soil temperature across 14 major biomes. Boxplots show the distributions of R for GloSVeT (orange), JRA-3Q (blue), and GLDAS (dark gray). Stars denote the mean values. T/ST, MBF, DBF, CF, Temp, BMF, F/T, GSS, GS, Med, FWS, and XS, are the abbreviations of 'Tropical/subtropical', '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.

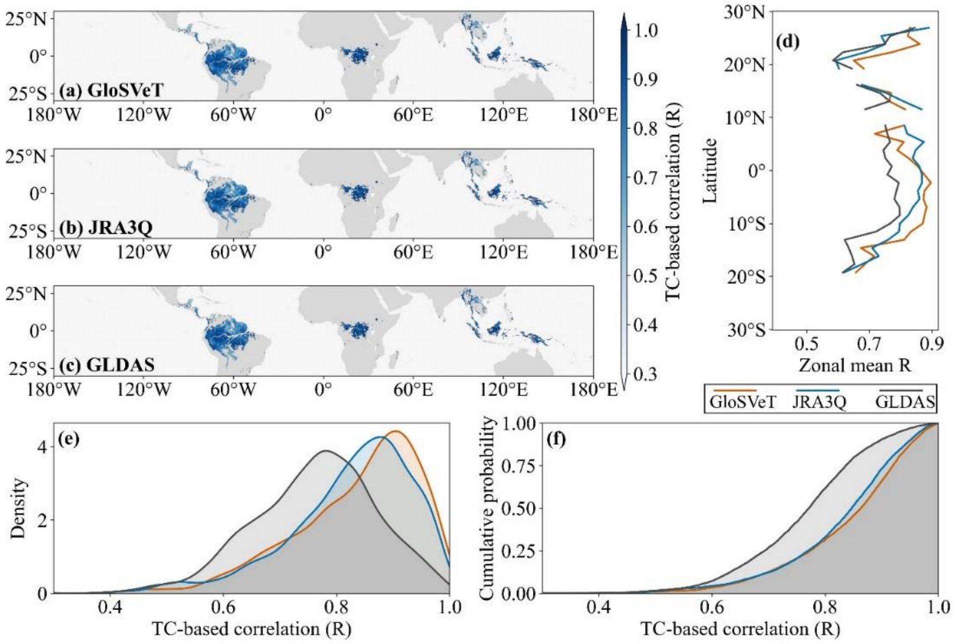

Figure 6: TC-based correlation (R) for vegetation temperature. Panels (a–c) illustrate the spatial distributions for GloSVeT, JRA-3Q, and GLDAS, respectively, with light gray and medium gray indicating the ocean and land backgrounds; (d) shows the zonal-mean R as a function of latitude; and panels (e)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

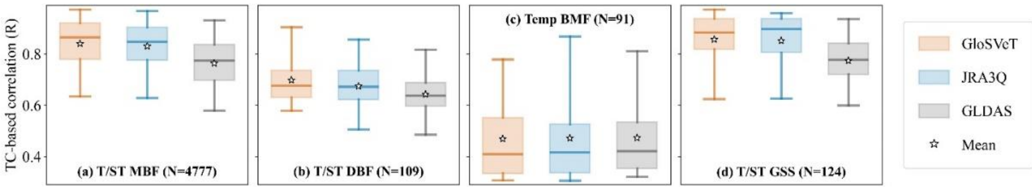

Figure 7: TC-based correlation (R) for vegetation temperature across four major biomes. Boxplots show the distributions of R for GloSVeT (orange), JRA-3Q (blue), and GLDAS (dark gray). Stars denote the mean values. T/ST, MBF, DBF, Temp, BMF, and GSS, are the abbreviations of 'Tropical/subtropical', 'moist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

Q2. How are the total percentages of statistically significant pixels calculated in section 4.3, paragraphs 2 & 3 (L366-390)? To me it seems like simple additions of all the percentages from each lag within each biome (shown in Figure 7). This wouldn't account for the fact that the grid is considered once for each lag (so the total is out of 6 * grid-size) and would lead to inflated percentages. If it's simple addition this needs to be corrected by the authors. If they use another method, and it's coincidental that that percentages add-up, this may need to be stated by the authors.

Reply:

Thank you for this detailed comment. We apologize for the ambiguity caused by our unclear description. In fact, for each valid pixel, correlations were first calculated under seven lag conditions from −3 to +3 months. Among the statistically significant lag correlations, only the lag with the largest absolute correlation coefficient was identified as the best lag. If more than one significant lag was detected, the best lag was determined only when its absolute correlation exceeded the second largest absolute correlation by at least 0.05. Therefore, each pixel had at most one retained result, namely the identified best lag and its corresponding correlation coefficient. The percentages shown in Figure 7 were then calculated by grouping these retained pixel-level results by lag month within each biome. Thus, the percentages across lag months can be summed, but they do not involve repeated counting of the same pixel under different lag conditions.

In the revised manuscript, we clarified this procedure in the Methods section and revised the relevant Results text and the caption of Figure 7 to avoid further ambiguity. The main revisions are listed below.

For the soil component, correlations were calculated between soil temperature anomalies and soil moisture anomalies for seven monthly lag conditions, namely −3, −2, −1, 0, +1, +2, and +3 months. Positive lags were defined as soil temperature anomalies preceding soil moisture anomalies, whereas negative lags indicate the opposite temporal ordering. These lagged correlations were used to identify the monthly window in which the two variables showed the strongest association, rather than to estimate an exact physical response time. For each valid pixel, only lagged correlations passing the significance test ($p < 0.05$) were retained as candidates. If only one candidate lag was identified, it was selected directly. If multiple candidate lags were identified, the lag with the largest absolute correlation coefficient was selected as the best lag only when its absolute correlation exceeded the second largest absolute correlation by at least 0.05. The retained best lag and its corresponding correlation coefficient were then used in the subsequent analyses. (Lines 310-319)

Across biome types (Figures 8b–o), the proportion of pixels with an identified best lag and the corresponding strength of soil temperature–moisture correlations exhibit distinct ecological patterns. (Lines 480-481)

Figure 8: Correlation between GloSVeT soil temperature anomalies and ESA CCI soil moisture anomalies for the best lag. Panel (a) shows the spatial distribution of correlations for pixels with an identified best lag, where red and blue indicate positive and negative relationships, respectively. Panels (b–o) present the distributions of these correlations across 14 major biomes. Boxplots with different colors denote the identified lag months from −3 to +3. The numbers above each box indicate the percentage of pixels assigned to each identified lag month within the corresponding biome, with darker colors and larger font sizes representing higher proportions. N denotes the total number of pixels in each biome category. (Lines 515-520)

Q3. L280-290 attributes a lot of the error in soil temperature to vegetation cover, and vice versa. If this is the case are the temperatures truly separable? This contradictory to the study's aims, but I am not an expert in FuSVeT.

Reply:

Thank you for this thoughtful comment. We agree that our original wording may have caused confusion. Our intention was not to suggest that soil and vegetation temperatures are inseparable, but to point out that the retrieval uncertainty of a component can increase when that component is weakly represented in the mixed-pixel thermal signal. In FuSVeT, soil and vegetation temperatures are separated using multi-temporal LST observations together with FVC constraints. Therefore, the two components are separable in principle, but their retrieval accuracy depends on how strongly each component contributes to the mixed-pixel signal. For example, under dense vegetation cover in summer, the soil component is relatively weakly constrained, which may increase soil temperature uncertainty. Conversely, under sparse canopy conditions in winter, vegetation temperature is less strongly constrained, leading to larger retrieval errors. This reflects an expected uncertainty pattern in component temperature separation rather than a contradiction to the separability assumption.

To avoid this ambiguity, we revised the corresponding text in the Results section. The revised sentence is as follows: "For soil temperature, performance was slightly lower in summer (R^2= 0.84, RMSE = 2.2 K), likely due to the weaker constraint on the soil component under dense vegetation cover (Liu et al., 2020b; Zhan et al., 2011; Zhan et al., 2013). In contrast, vegetation temperature retrievals showed larger errors in winter (RMSE = 3.0 K), which may be attributed to the reduced contribution of the vegetation component to the mixed pixel thermal signal under sparse canopy conditions." (Lines 366-370)

Minor Considerations

Q4. L50-52: If there is strong coupling between air temperature and canopy temperature.

Reply:

Yes, the strong coupling between air temperature and canopy/foliage temperature is supported by previous studies. For example, Zhan et al. (2011) stated that "*Foliage temperatures are closely connected to the soil surface and air temperatures through the interactions between the canopy elements and the environment [14], [49]. Being easily acquired from local weather stations, the air temperatures have been demonstrated as approximations for the leaf temperatures in most cases [50]. Although the foliage temperatures are strongly influenced by soil water depletion and some meteorological factors, the foliage temperatures have been used to explain and validate the accuracies of the inverted vegetation temperatures in previous studies [10].*"

Q5. L84-85: Mentioned the generalized three-cornered hat but not used in this paper. My suggestion would be to remove this and also briefly expand on the triple collocation method.

Reply:

Thank you for this helpful suggestion. We have removed the description of TCH to avoid unnecessary confusion, since it was not used in this study. We also briefly expanded the explanation of TC in the Introduction.

The revised text is as follows: "Among them, the triple collocation (TC) provides a practical tool for quantifying the relative errors and mutual consistency among three independent datasets that observe the same target variable (Gruber et al., 2016; Park et al., 2023; Stoffelen, 1998; Wei et al., 2024). Although TC was originally designed for other geophysical variables such as soil moisture and precipitation, its demonstrated robustness and scalability make it particularly promising for evaluating component temperature products, offering a pathway toward large-scale, physically consistent assessments of spatiotemporal stability, internal coherence, and overall applicability." (Lines 97-103)

Q6. Table 1: This shows that skin temperature is the monthly mean but does not clarify the measure used for the other datasets. Add whether other datasets are means or not.

Reply:

Thank you for this helpful comment. We have revised Table 1 to clarify the temporal resolution used for each dataset as follows. (Line 123)

Variable	Dataset	Spatial Resolution	Temporal Resolution	Period	Purpose
Land surface temperature	MOD11C3/ MYD11C3	0.05 °	Monthly mean at four overpass times	2003–2023	Input thermal observations for constructing the separation model
Skin temperature	ERA5-Land	0.10 °	Monthly mean at 24 hourly time steps	2003–2023	Model parameter simplification and constraint
Fractional vegetation cover	GLASS FVC	0.05 °	8-day composite	2003–2023	Pixel-scale structural parameterization for decomposition
Land cover	MCD12C1	0.05 °	Yearly	2003-2023	Masking of non-vegetated areas (water bodies and snow/ ice)

Table 1. Data used for the generation of the GloSVeT product

Q7. 112: is MODIS suitable post-2021? I believe either one or both of the satellites are in orbital drift. Is this important to the GloSVeT development?

Reply:

Thank you for this important comment. We agree that the orbital drift of Terra and Aqua after the end of regular orbit-maintenance maneuvers may affect the actual local overpass times of MODIS observations, especially after 2021. This issue is relevant to GloSVeT because MODIS LST provides the primary satellite thermal constraint in the FuSVeT framework.

However, the orbital drift does not invalidate the use of MODIS LST in GloSVeT. First, GloSVeT is generated at the monthly scale, and MODIS monthly LST provides aggregated thermal information rather than individual instantaneous observations. Second, the FuSVeT framework does not rely solely on MODIS overpass-time observations to reconstruct the full diurnal cycle. ERA5-Land skin temperature is used to provide continuous diurnal-cycle information and to constrain the temperature-independent parameters of the MDC model, while MODIS LST is mainly used to adjust the temperature level of the retrieved soil and vegetation components. Therefore, the potential influence of MODIS orbital drift is partly mitigated by the multi-source framework.

Nevertheless, we agree that the drift in actual MODIS overpass time after 2021 may introduce additional temporal sampling uncertainty, particularly for the last years of the record and for analyses that are sensitive to small long-term trends. We have therefore added this issue as a limitation in the revised manuscript and clarified that future updates of GloSVeT should further incorporate orbit drift correction or use additional sensors such as VIIRS.

The revisions are as follows: "A related input data issue is the orbital drift of the Terra and Aqua satellites during the later MODIS period, especially after 2021, which may introduce temporal sampling uncertainty because the actual overpass times deviate from the nominal MODIS overpass times. Although this effect is partly mitigated by monthly scale processing and the use of ERA5-Land diurnal cycle information, it may still influence the results toward the end of the data record. This issue could be further addressed by incorporating orbit drift correction (Julien and Sobrino, 2025; Liu et al., 2019) or by introducing additional sensors, such as VIIRS, within the FuSVeT framework." (Lines 678-684)

Q8. L117-118: How are 4 instantaneous observations per month used to construct the mean MDC? Are not more observations needed produce a mean MDC?

Reply:

Thank you for this comment. In the revised manuscript, we have clarified this description to avoid ambiguity, as follows: "These observations, together with the continuous diurnal cycle information derived from ERA5-Land hourly skin temperature, are used to construct the monthly mean diurnal cycle (MDC) model, which forms the basis for separating soil and vegetation component temperatures." (Lines 131-133)

Q9. L138-140: What do you do about urban pixels?

Reply:

Thank you for this question. Urban and built-up pixels were retained in GloSVeT, mainly for spatial completeness and because such pixels may still contain mixed vegetation and non-vegetation background signals at the 0.05 °scale. However, we agree that the physical interpretation of the soil component over urban areas is different from that over natural soil surfaces.

In the revised manuscript, we clarified the treatment of urban pixels in the section 2.1: By contrast, urban and built-up pixels (IGBP type = 13) were retained for spatial completeness, since such pixels may contain mixed vegetation and non-vegetation background signals at the 0.05 °scale. (Lines 154-156)

We also explicitly discussed this limitation in Section 5.3: However, over complex surfaces such as urban or mountainous regions, the so-called soil component temperature may include impervious surfaces, exposed rocks, or other non-vegetation materials, and should therefore be interpreted as a non-vegetation background component temperature in these areas … Consequently, GloSVeT is more suitable for characterizing the surface radiative thermal states of visible soil/background and vegetation components. For applications involving under-canopy soil temperature, sub-snow soil temperature, urban background thermal states, or vertical heat transfer processes, GloSVeT should be interpreted cautiously in combination with land cover, snow cover, topographic, or canopy structural information. (Lines 655-665)

Q10. L195: The reference here is Göttsche & Olesen 2009 and the text states the "well established GOT09 model", though this is not in the reference. There is a "Goe2008" model, is this what the authors are referring to?

Reply:

We thank the reviewer for pointing this out. Yes, the model referred to here is the Goe2008 model proposed by Göttsche and Olesen (2009). We used the term "GOT09" because this name has been widely adopted in subsequent evaluations and applications of DTC models. For example, Duan et al. (2012) and Hong et al. (2018) referred to the Goe2008 model as GOT09. Therefore, in this manuscript, we retained the term "GOT09" to be consistent with the terminology commonly used in later DTC model comparison studies. We apologize for any confusion caused by this nomenclature.

Duan, S.-B., Li, Z.-L., Wang, N., Wu, H., & Tang, B.-H. (2012). Evaluation of six land-surface diurnal temperature cycle models using clear-sky in situ and satellite data. *Remote Sensing of Environment*, 124, 15-25.

Hong, F., Zhan, W., Göttsche, F.-M., Liu, Z., Zhou, J., Huang, F., Lai, J., & Li, M. (2018). Comprehensive assessment of four-parameter diurnal land surface temperature cycle models under clear-sky. *ISPRS Journal of Photogrammetry and Remote Sensing*, 142, 190-204.

Q11. L219: State explicitly which parameters are temperature-independent and -dependent. Will make the reader do less work and improve ease of reading.

Reply:

Done.

Q12. L230-231: In the current context, this maybe a slightly confusing/misleading of the word "downscaled" as the values are just being duplicated. Suggest finding a different description.

Reply:

Thanks for this helpful suggestion. In revised manuscript, we changed this sentence to "The retrieved parameters were subsequently mapped to the 0.05° grid by uniformly assigning each 0.10° estimate to its corresponding 2 × 2 finer pixels." (Lines 257-258)

Q13. L246: When "corresponding daytime values [are] shifted by 12 hours", is this using the MDC? If not, how?

Reply:

We apologize for the unclear expression. In the revised manuscript, we clarified that this correction was applied to the nighttime viewing time layer rather than to the LST values. The revised text is as follows: "a correction was applied to the nighttime viewing time layer. Specifically, when the nighttime viewing time exceeded its theoretical range of 00:00–03:00, it was replaced by the corresponding daytime viewing time shifted by 12 hours under the 24 hour clock. This correction was applied only to the viewing time information and did not modify the MYD11C3 LST values." (Lines 284-287)

Q14. L275-276: Offsets partially attributed to the "imperfect representation of near sunrise temperature conditions" – but isn't this what is resolved in Göttsche & Olesen 2009?

Reply:

Thank you for pointing this out. We agree that the Göttsche and Olesen model was developed to improve the representation of diurnal temperature dynamics, especially for the near-sunrise period. However, the issue discussed here is not a limitation of the model formulation itself, but rather the limited temporal sampling of the satellite observations used to constrain the fitted monthly mean diurnal cycle. In this study, only four MODIS overpass-time observations were available, and none of them directly sampled the near-sunrise period. Therefore, even with an improved diurnal temperature cycle model, the fitted curve still be insufficiently constrained near sunrise (Liu et al., 2025). To avoid this ambiguity, we revised the sentence and removed the citation to Göttsche and Olesen (2009) from this specific explanation. The revised sentence is: "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. (Lines 339–343)" and "Similar to the result of internal closure check, this systematic offset may result from the uneven temporal distribution of satellite observations, especially the lack of direct constraints near sunrise, which can lead to insufficient representation of near-sunrise temperature conditions in the fitted monthly mean diurnal cycle (Liu et al., 2025). (Lines 357–360)"

Q15. L283-285: Reduced correlation in summer Ts partially attributed to fewer stations, though Autumn does well with only ~30 more stations. Similarly, with Tv in winter, as spring bias is smaller with on an extra 11 stations. Are these dominated by the vegetation cover? Suggest is more important.

Reply:

Thank you for this helpful comment. We agree that the seasonal differences in validation accuracy should not be mainly attributed to the number of available observations. In the revised manuscript, we removed this explanation and clarified that the lower summer performance for soil temperature and the larger winter error for vegetation temperature are more likely related to seasonal changes in vegetation cover and canopy conditions, which affect the contribution of each component to the mixed pixel thermal signal and hence the retrieval constraint. The revised text is as follows: "For soil temperature, performance was slightly lower in summer (R^2= 0.84, RMSE = 2.2 K), likely due to the weaker constraint on the soil component under dense vegetation cover (Liu et al., 2020b; Zhan et al., 2011; Zhan et al., 2013). In contrast, vegetation temperature retrievals showed larger errors in winter (RMSE = 3.0 K), which may be attributed to the reduced contribution of the vegetation component to the mixed pixel thermal signal under sparse canopy conditions." (Lines 366-370).

Q16. L289: Could be worth a discussion about the RMSE values, as GCOS requirement for LST is <1K if I recall. How does this impact the usefulness of the dataset?

Reply:

Thank you for this helpful comment. We agree that the <1 K benchmark is an important reference for standard LST products. However, GloSVeT provides monthly mean soil and vegetation component temperatures separated from mixed pixel LST, rather than direct LST retrievals. Therefore, its RMSE values also include uncertainties associated with component separation and validation representativeness. To avoid overstatement, we revised the original wording and described the RMSE values more neutrally. The revised text is as follows: "Overall, these results indicate that GloSVeT achieves consistent and physically reasonable performance across seasons, with strong correlations ($R^2 > 0.9$ in most cases) and RMSE values typically below 2 K, effectively reproducing surface component temperatures at the monthly scale.” (Lines 380-382)

Q17. Figure 4: Minor comment – change “vermillion” to orange/red/pink, for ease of reading.

Reply:

Thanks for this suggestion. We have changed “vermillion” to “orange” in the revised figure caption for easier reading.

Q18. Figure 6: Personally, I prefer more explanatory figure captions, over “Same as…but….”, so would suggest repeating the figure 4 caption and mentioning this it’s for vegetation temperature. There’s only 4 biomes, so these could be repeated too.

Reply:

Thank you for your careful suggestion. In the revised manuscript, we have made the caption of Figure 6 more self-contained and explanatory. The revised caption is: “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.” (Lines 460-464)

In addition, we also revised the caption of Figure 5 (now Figure 6) to avoid a simplified “same as” description. The revised caption is: “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.” (Lines 454-458)

Q19. L356-357/Figure 7a: Is this in reference to the 0 lag?

Reply:

Thank you for this detailed comment. Figure 7a (Figure 8a in the new manuscript) does not refer specifically to the 0 month lag. As clarified in our response to Q2, for each pixel, correlations were calculated under seven lag conditions from −3 to +3 months, and only the identified best lag and its corresponding correlation coefficient were used in the subsequent analyses. Therefore, Figure 8a shows the spatial distribution of correlations for pixels with an identified best lag, rather than the correlation pattern at 0 month lag. We have clarified this point in the revised manuscript and in the caption of Figure 8.

Q20. Figure 8: The changing font size/colour is quite hard read in some instances and may be less accessible to some. I would suggest introducing a twin-axis or another subplot (e.g. barplot) to demonstrate the numbers and proportions of pixels.

Reply:

Thank you for this helpful suggestion. We have revised Figure 8 (Figure 9 in the new version) to improve readability and accessibility. Instead of using changing font sizes and colours to indicate the proportion of significant pixels, we added a separate bar plot as panel (c). The boxplots in panel (b) now focus on the distributions of significant correlations, while panel (c) directly shows the proportion of significant pixels within each biome, together with the total number of pixels. The revised Figure is as follows. (Lines 544-549)

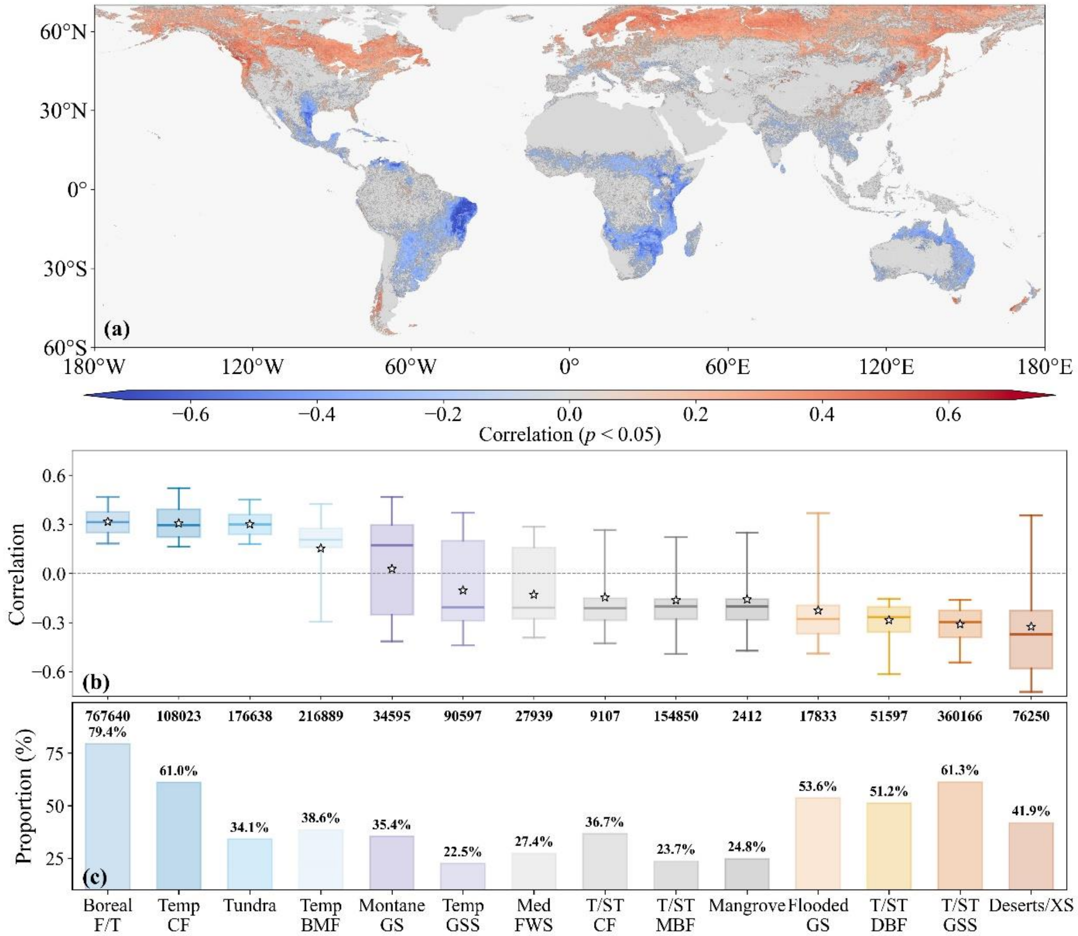

Figure 9: Correlation between GloSVeT vegetation temperature anomalies and GOSIF SIF anomalies. Panel (a) shows the spatial distribution of significant correlations (p < 0.05), where red and blue indicate positive and negative relationships, respectively. Panel (b) presents the distribution of significant correlations across 14 major biomes. Panel (c) summarizes the proportion of significant pixels within each biome. In panel (c), the numbers above the bars indicate the total number of pixels in each biome category, and the percentages indicate the corresponding proportion of significant pixels.

Q21. L430-431: What are the uncertainties associated with these trends?

Reply:

Done. We have added the standard errors and significance levels of the estimated trends, such that: "with slopes of 0.44 ± 0.04 K/decade for soil temperature and 0.39 ± 0.04 K/decade for vegetation temperature (both $p < 0.001$)." (Lines 556 to 557)

Q22. Figure 9: It could be suggested that there are two regimes in this timeseries, pre- and post-2015. Are there any changes in the input data between 2013-2016 (e.g. satellite sensor changes)?

Looks like a possible step change, which could impact trends – worth confirming if a change is present to be sure they are real trends.

Reply:

Thank you for this careful observation. GloSVeT was generated using a consistent set of inputs, including MOD11C3/MYD11C3, ERA5-Land, GLASS FVC, and MCD12C1, throughout 2003–2023. We also checked the valid pixel counts and found no abrupt change around 2015. Therefore, the apparent higher anomalies after 2015 are unlikely to result from a step change in the input data or processing workflow. Instead, they are more likely related to strong climate variability during this period, especially the 2015–2016 El Niño event, which was reported as one of the strongest El Niño events on record according to the NOAA Niño 3.4 index record (https://psl.noaa.gov/data/timeseries/month/DS/Nino34_CPC/).

Q23. L448: What do you mean by "major warming and cooling centres"?

Reply:

In the revised manuscript, we corrected this wording as "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." (Lines 573 to 574)

Q24. L453: Suggest swapping "patchier" for "more heterogeneous".

Reply:

Thank you for this suggestion. We have replaced "patchier" with "more heterogeneous" in the revised manuscript (Line 580).

Q25. L453: What impacts would "limited land area and weaker radiative forcing contrast" have on per grid-cell temporal trends? Is this related to uncertainty?

Reply:

Thank you for this comment. To avoid ambiguity, we corrected the text as "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." (Lines 580-581)

Q26. Figure 11: Personally, I prefer more explanatory figure captions, over "Same as…but….", so would suggest repeating the figure 10 caption and mentioning this it's for July.

Reply:

Done. In the revised manuscript, we revised the caption as "Figure 12: Linear trends of GloSVeT component temperatures in July for (a) soil and (b) vegetation during 2003–2023. Black dots indicate pixels with statistically significant trends ($p < 0.05$)." (Lines 600-601)

Q27. L522: Is this statistically significant, or significant in magnitude/importance?

Reply:

In the revised manuscript, we replaced 'significant' with 'pronounced' to avoid misunderstanding (Line 734).