Response to Reviewer#3 Comments

We sincerely thank the reviewers for their thorough and insightful evaluation of our manuscript. Their detailed comments and constructive suggestions have been invaluable in improving the quality of this work. We are grateful for the considerable time and expertise they have dedicated to this review. In the following, we address each comment individually. The reviewer's comments are in black, and our responses are provided in blue.

Comment 1:

Differences in soil depth and vertical discretization can substantially affect heat and water transport in soils. Given that the three systems adopt different soil layering schemes, it remains unclear to what extent the simulated differences in soil thermal and hydrological processes arise from physical process representations versus structural differences in vertical discretization.

Response:

We thank the reviewer for this important and perceptive comment. We agree that differences in soil depth and vertical discretization can substantially influence the simulation of soil thermal and hydrological processes, and that this introduces inherent challenges in attributing inter-model differences to either physical process representations or structural configuration. Since JSBACH and CLM differ simultaneously in their vertical discretization, soil column depth, and physical parameterizations, separating the contributions of these factors in a fully controlled manner would require targeted sensitivity experiments with consistent model physics and varying vertical configurations. While such experiments would be scientifically valuable, they are beyond the scope of the present study.

This is a well-recognised challenge in the land surface modelling community. Koven et al. (2013) and Slater and Lawrence (2013) demonstrated a broad range of inter-model spread in simulated soil temperatures across CMIP5 models, attributing the uncertainty to a combination of snow insulation, land–atmosphere coupling, and vertical model resolution. González-Rouco et al. (2021) further demonstrated that increasing JSBACH's soil column depth from 9.83 m to over 1400 m substantially reduced deep soil warming by 0.5-1.5 K in scenario simulations, with energy storage 3-5 times larger in the deeper configuration, underscoring that vertical structure alone can introduce biases independent of process parameterizations.

The primary goal of this study is to evaluate how two widely used land surface models within ESMs, with fundamentally different configurations: JSBACH employing a comparatively coarse vertical discretization (5 soil layers to ∼9.83 m depth) and CLM5 a much finer soil column structure (20 soil layers extending to 8.6 m depth, plus 5 bedrock layers to ∼50 m), represent seasonally frozen ground characteristics. This comparison provides direct insight into their respective strengths and limitations for frozen ground applications.

Therefore, in response to the reviewer's comment, we agree that this distinction deserves explicit acknowledgement in the manuscript. **Accordingly, we will add the following paragraph in the revised manuscript at Line 324:**

In addition, differences in vertical soil discretization can affect heat and water transport in the soil column. Koven et al. (2013) and Slater and Lawrence (2013) identified vertical resolution, together with snow insulation and land surface physics, as key sources of uncertainty in simulating cold-region soil thermal states, while González-Rouco et al. (2021) demonstrated that soil column depth alone can substantially influence long-term subsurface temperature dynamics. Therefore, the inter-model differences reported in this study should be interpreted as reflecting the combined influence of structural configuration and physical process representation, rather than being attributable to either factor in isolation. Furthermore, since all simulations are driven by .... (Line 324-327)

We will also update Table 1 in the revised manuscript to include detailed information on the soil layer structures of both models as follows:

Table 1: Key structural and parametric differences in soil and snow process representations between JSBACH and CLM

Model features	**JSBACH**	**CLM**
Total soil column depth	∼9.8 m	∼49.5 m (bedrock below 8.6 m)
Number of soil layers (default)	5 soil layers	20 soil layers + 5 bedrock layers
Soil layer node depths (m)	0.0325, 0.192, 0.7755, 2.683, 6.984	0.01, 0.04, 0.09, 0.16, 0.26, 0.4, 0.58, 0.8, 1.06, 1.36, 1.7, 2.08, 2.5, 2.99, 3.58, 4.27, 5.06, 5.95, 6.94, 8.03, 9.795, 13.328, 19.483, 28.871, 41.998
Soil layer thickness (m)	0.065, 0.254, 0.913, 2.902, 5.7	0.02, 0.04, 0.06, 0.08, 0.12, 0.16, 0.2, 0.24, 0.28, 0.32, 0.36, 0.4, 0.44, 0.54, 0.64, 0.74, 0.84, 0.94, 1.04, 1.14, 2.39, 4.676, 7.635, 11.14, 15.115
Number of snow layers (default)	5 (hydrologically inactive)	10
Number of PFTs	11	15
Concept of supercooled water	Yes (Niu and Yang, 2006)	Yes (Niu and Yang, 2006)
Soil heat capacity scheme	(De Vries (1953))	(De Vries (1953))
Soil thermal conductivity	(Johansen (1977))	(Farouki (1981))
Bottom boundary condition	Zero flux	Zero flux
Snow density formulation	Temperature dependent	Multi-process: compaction, destructive and constructive, melt metamorphism (Anderson, 1976)
Fresh snow density parameterization	50 kg m^{-3} (constant)	Temperature dependent
Snow thermal conductivity	(Calonne et al. (2011))	(Jordan (1991))

Comment 2:

Soil ice plays an important role in regulating soil thermal properties and heat transfer. Although ERA5-Land does not explicitly represent soil ice, a more direct comparison between JSBACH and CLM in terms of soil ice processes could provide additional insight.

Response:

We thank the reviewer for this comment. The above mentioned aspect has already been addressed in the manuscript. While ERA5-Land does not explicitly represent soil ice, we provided a direct comparison of soil ice content between JSBACH and CLM in the Appendix (Figure B1), which offers additional insight into the role of soil ice processes. Furthermore, taking advantage of the availability of the ESA CCI surface soil moisture dataset, we have also included a small evaluation of simulated liquid soil moisture in the top layer against this reference data (Figure B2).
These figures together reveal important differences in soil ice and liquid water dynamics between the two models. In cold regions, CLM simulates a larger fraction of total soil moisture in the frozen/ice phase compared to JSBACH (Figure B1), indicative of stronger near-surface soil freezing in CLM. Evaluation of top-layer liquid soil moisture against ESA CCI observations (Figure B2) reveals that CLM tends to underestimate liquid soil moisture across boreal and Arctic regions. While summer thaw periods lead to increased liquid water content, autumn soil moisture in CLM remains systematically lower than observed, suggesting persistently drier near-surface conditions prior to freeze-up that may further promote and accelerate the onset of freezing in autumn.
The corresponding information is already included in the manuscript **(Lines 276-280)**.

Comment 3:

While JSBACH appears to perform better in simulating SFG, it lacks an explicit coupling between snow and soil water processes. This missing process may be important, particularly as the manuscript emphasizes the role of snow. Beyond albedo and insulation effects, snowmelt dynamics at high latitudes may also significantly influence soil thermal regimes. There is a concern that compensating errors among missing or simplified processes could lead to apparently improved performance. These limitations should be more explicitly acknowledged, potentially in the Conclusions.

Response:

We would like to clarify that in the manuscript and conclusion, we have only stated that JSBACH better captures frozen ground extent compared to the other two datasets. For other aspects, such as SFG and snow cover characteristics, JSBACH performs poorly due to different process representations (e.g., underestimation of snow depth and weaker insulation), just as CLM and ERA5L do for different reasons. We agree with the reviewer that, beyond albedo and insulation effects, snowmelt dynamics can play an important role in influencing soil thermal regimes, particularly in high-latitude regions. While our analysis discusses the role of snow cover timing (e.g., onset and end date), we acknowledge that the explicit role of snowmelt processes and their interaction with soil hydrology was not sufficiently emphasized. We also agree that this limitation should be more clearly acknowledged, particularly in the context of interpreting model performance.
Therefore, we will revise the manuscript (1. Lines 451-452 and 2. Lines 461-462) to acknowledge this limitation and highlight the importance of snowmelt dynamics.

1. At most stations, the models exhibited premature soil freezing and delayed spring thaw, reflecting persistent challenges in capturing autumn land-atmosphere interactions and snow evolution processes **(including snowmelt dynamics)**.
2. "Future studies should focus **on site-level performance to investigate the combined effects of snow accumulation and snowmelt processes**, vegetation, soil properties, and soil moisture on soil freeze-thaw dynamics, in order to enhance land surface model performance."

Comment 4:

In Sect. 3.3.4, the authors provide a detailed analysis of controlling factors. However, a schematic diagram illustrating how these factors influence soil temperature and, consequently, SFG would greatly improve clarity, especially given the extensive use of abbreviations.

Response:

We thank the reviewer for this important suggestion. In response, we have decided to add a partial correlation table, as also suggested by Reviewer #2, together with a schematic diagram summarizing the statistically significant partial correlations between SFG characteristics, snow cover properties, and seasonal thermal drivers across observations and models (Fig. 1). The controlling variables used for each variable pair are listed in Table 2 and were selected based on physically motivated relationships to isolate direct from indirect pathways. To further visualise the relationships presented in Table 3, including both pearson correlation and partial correlation values calculated using the controlling parameters listed in Table 2, we constructed a schematic diagram (Fig. 1) for observations and each model. This diagram helps distinguish genuine physical drivers from spurious inter-correlations.

Variable definitions (for reference):

- SFG_Onset - onset date of seasonally frozen ground at 20 cm depth.
- SFG_End - end date of seasonally frozen ground at 20 cm depth
- SFGD - seasonally frozen ground duration (SFG_End - SFG_Onset)
- SC_Onset - snow cover onset date
- SC_End - snow cover end date
- SCD - snow cover duration (SC_End - SC_Onset)
- Autumn_AT - autumn air temperature
- Winter_AT - winter air temperature
- Spring_AT - spring air temperature
- Winter_SD - winter snow depth

Partial correlations (pc) were computed to remove the influence of shared drivers, particularly Autumn_AT, Spring_AT, and Winter_SD, which co-vary with multiple target variables simultaneously. This approach allows us to determine, for example, whether the relationship between SFG_Onset and SC_Onset is direct or primarily inherited from a common autumn temperature signal.

To illustrate the calculation, we provide an example of the partial correlation between parameters A and B, with C held as the control variable:

$$pc_{AB\cdot C} = \frac{R_{AB} - R_{AC}R_{BC}}{\sqrt{(1 - R_{AC}^2)\,(1 - R_{BC}^2)}} \tag{1}$$

Observation: In the observations, SC_Onset and Autumn_AT show a direct moderate positive relationship (R = 0.3), indicating that warmer autumn temperatures delay snow cover onset. SC_Onset and SFG_Onset exhibits virtually no direct relationship (partial r = 0.02 after controlling for Autumn_AT), indicating that frozen ground onset is not directly driven by snow arrival and is instead likely governed by local factors such as soil properties, vegetation cover, soil moisture, and microclimatic conditions. Similarly, after controlling for SC_Onset, the relationship between Autumn_AT and SFG_Onset is very weak (pc = 0.14), suggesting only limited direct influence of Autumn_AT on soil freezing onset. SC duration is primarily controlled by Winter_SD (pc = -0.52), Spring_AT (pc = -0.50), and Autumn_AT (pc = -0.33). SC_End shows strong dependence on Spring_AT (pc = -0.66) and Winter_SD (pc = +0.52), and a moderate relationship with SFG_End after controlling for spring temperature (pc = +0.37). In contrast, SFG_End is only weakly directly influenced by Spring_AT once the effect of SC_End is removed (pc = -0.20), highlighting that soil thaw timing is mediated through snow cover dynamics rather than temperature alone. In the observations, SFG duration shows a moderate negative relationship with both Autumn_AT (pc = -0.37) and Spring_AT (pc = -0.21) after controlling for SC duration, indicating that colder autumn and spring conditions contribute weakly to longer frozen ground duration independent of snow cover persistence.

ERA5L: In ERA5L, the schematic indicates no direct relationship between SFG_Onset and Autumn_AT. SC_End, however, shows relatively strong relationships with both SFG_End (pc = 0.85) and Spring_AT (pc = -0.69) when the influence of other variables is accounted for. In addition, the direct influence of Spring_AT on SFG_End (pc = -0.42) remains stronger than observed, suggesting that ERA5L overestimates the sensitivity of soil thaw to spring warming. A higher SFG duration-SC duration correlation (R= 0.53) in ERA5L than in observations (R = 0.38)) indicates that the model overestimates the coupling between snow cover duration and frozen ground duration. SFG duration maintains moderate negative relationships with both Autumn_AT (pc = -0.32) and Spring_AT (pc = -0.36) after controlling for SC duration, indicating a slightly stronger sensitivity of frozen ground duration to seasonal warming than observed.

CLM: CLM reproduces the correct directional sign across all major linkages, but SC_Onset's dependence on Autumn_AT is amplified (R = 0.73), suggesting snow onset timing is highly sensitive to autumn temperature conditions. Most critically, the SC_Onset-SFG_Onset relationship collapses to near zero, while the Autumn_AT-SFG_Onset relationship remains weak (pc = +0.09), which may suggest that the relatively high snow depth in CLM could partially thermally isolate the soil, reducing the influence of snow onset timing and autumn air temperature on soil freezing at 20 cm depth. SFG duration shows a stronger dependence on SC duration (R = 0.57) while SC duration is strongly controlled by Winter_SD (pc = +0.69), Spring_AT (pc = -0.52) and Autumn_AT (pc = -0.35), reflecting the dominant role of snow accumulation and melt-season temperature in regulating snow cover duration. Furthermore, the SC_End-SFG_End partial correlation is substantially amplified (pc = 0.84) after controlling for Spring_AT, indicating that snow disappearance strongly regulates soil thaw timing in CLM. Once SC duration and Spring_AT are controlled, Autumn_AT's independent effect on SFG duration weakens to pc = -0.20 in CLM (vs -0.37 in observations), while Spring_AT's independent effect (pc = -0.33) remains slightly stronger than observed (pc = -0.21), suggesting that CLM further suppresses the direct autumn thermal pathway to the soil while retaining an elevated spring signal.

JSBACH: JSBACH exhibits a distinctly different relationship structure, likely reflecting its substantial snow depth underestimation. Unlike observations and CLM, SC_Onset shows virtually no dependence on Autumn_AT (R = 0.07), suggesting that snow accumulation timing in JSBACH is largely disconnected from autumn thermal conditions. The SFG_Onset and Autumn_AT partial correlation surges to pc = 0.71, the strongest signal in the entire network and far exceeding the observed value (pc = 0.14), while SC_Onset and SFG_Onset show no meaningful relationship (pc = -0.04), together confirming that in the absence of adequate snow cover, Autumn_AT drives soil freezing directly without the moderating influence of snow. SFG duration shows a stronger dependence on SC duration (R = +0.56) while SC duration is strongly controlled by Winter_SD (pr = +0.66), Spring_AT (pc = -0.35), and Autumn_AT (pc = -0.27). The SC_End and SFG_End coupling weakens substantially (pc = 0.18 vs 0.37 in observations), while Spring_AT's independent effect on SFG_End remains quite strong even after controlling for SC_End (pc = -0.68), indicating that without adequate snow insulation, spring air temperature forces soil thaw directly rather than through snow cover. In JSBACH, Autumn_AT's independent effect on SFG duration (pc = -0.52) and Spring_AT's independent effect (pc = -0.54) both remain considerably stronger than observed, confirming that seasonal air temperatures control frozen ground duration directly, bypassing snow insulation. In JSBACH, the sign change for SFG_Onset (strong negative dependency on SFGD, R = -0.77) together with the positive dependency of SFG_End on SFG duration (R = +0.84) suggests that both early onset and delayed end contribute to longer frozen ground duration, whereas in other datasets SFG duration shows stronger control by SFG_End.

Table 2: Relationships between seasonal air temperatures, snow depth, snow cover, and seasonally frozen ground characteristics with controlling factors and physical interpretation.

Parameters	Controlling Factors	Remarks
Autumn_AT × Winter_SD	–	Independent environmental drivers
Winter_AT × Winter_SD	–	Independent environmental drivers
SC_Onset × Autumn_AT	–	Direct autumn temperature control on snow onset.
SFG_Onset × SC_Onset	Autumn_AT	Remove shared effect of autumn temperature.
SFG_Onset × Autumn_AT	SC_Onset	Isolate direct thermal effect on soil freezing.
SCD × Autumn_AT	Winter_AT, Spring_AT, Winter_SD	SCD influenced by seasonal AT and snow depth; control removes indirect pathways, isolating direct effect.
SCD × Spring_AT	Winter_AT, Autumn_AT, Winter_SD	
SCD × Winter_SD	Winter_AT, Spring_AT, Autumn_AT	
SC_End × Spring_AT	Winter_AT, Winter_SD	Isolate spring warming effect on snow disappearance.
SC_End × Winter_SD	Winter_AT, Spring_AT	Snow depth effect on melt date, excluding temperature & thaw.
SC_End × SFG_End	Spring_AT	Remove shared effect of spring temperature.
SC_End × SCD	–	Mathematically linked (duration = end - onset): no control needed.
SC_Onset × SCD	–	
SFG_End × Spring_AT	SC_End	Direct spring temperature control on SFG onset.
SFGD × SCD	–	Includes the influence of all other factors to show total coupling effect.
SFGD × Autumn_AT	SCD, Spring_AT	Isolates the direct influence of autumn air temperature
SFGD × Spring_AT	SCD, Autumn_AT	Isolates the direct influence of spring air temperature
SFD_End × SFGD	–	Mathematically linked (duration = end - onset): no control needed.
SFG_Onset × SFGD	–	Mathematically linked (duration = end - onset): no control needed.

Table 3: Statistically significant (p value < 0.01) pearson and partial correlations between seasonal air temperatures, winter snow depth, snow cover, and seasonally frozen ground characteristics.

Parameters	**Pearson Correlation (R)**				**Partial Correlation (pc)**			
	Obs	**ERA5L**	**CLM**	**JSBACH**	**Obs**	**ERA5L**	**CLM**	**JSBACH**
Autumn_AT × Winter_SD	-0.59	-0.59	-0.51	-0.58	-	-	-	-
Winter_AT × Winter_SD	-0.34	-0.35	-0.16	-0.53	-	-	-	-
SC_Onset × Autumn_AT	0.3	0.71	0.73	0.07	-	-	-	-
SFG_Onset × SC_Onset	0.07	0.2	0.1	0.02	0.02	0.15	0.004	-0.04
SFG_Onset × Autumn_AT	0.15	0.13	0.13	0.71	0.14	-0.02	0.09	0.71
SCD × Autumn_AT	-0.7	-0.74	-0.68	-0.64	-0.33	-0.4	-0.35	-0.27
SCD × Spring_AT	-0.68	-0.7	-0.68	-0.52	-0.5	-0.5	-0.52	-0.35
SCD × Winter_SD	0.71	0.77	0.78	0.76	0.52	0.62	0.69	0.66
SC_End × Spring_AT	-0.74	-0.77	-0.72	-0.57	-0.66	-0.69	-0.65	-0.51
SC_End × Winter_SD	0.62	0.7	0.73	0.66	0.52	0.61	0.67	0.61
SC_End × SCD	0.89	0.91	0.92	0.84	-	-	-	-
SC_Onset × SCD	-0.34	-0.87	-0.87	0.02	-	-	-	-
SFG_End × Spring_AT	-0.57	-0.82	-0.72	-0.78	-0.2	-0.42	-0.23	-0.68
SFGD × SCD	0.38	0.53	0.57	0.56	-	-	-	-
SFGD× Autumn_AT	-0.52	-0.58	-0.53	-0.73	-0.37	-0.32	-0.2	-0.52
SFGD × Spring_AT	-0.42	-0.59	-0.6	-0.71	-0.21	-0.36	-0.33	-0.54
SFG_End × SFGD	0.52	0.69	0.75	0.84	-	-	-	-
SFG_Onset × SFGD	0.26	0.31	0.14	-0.77	-	-	-	-

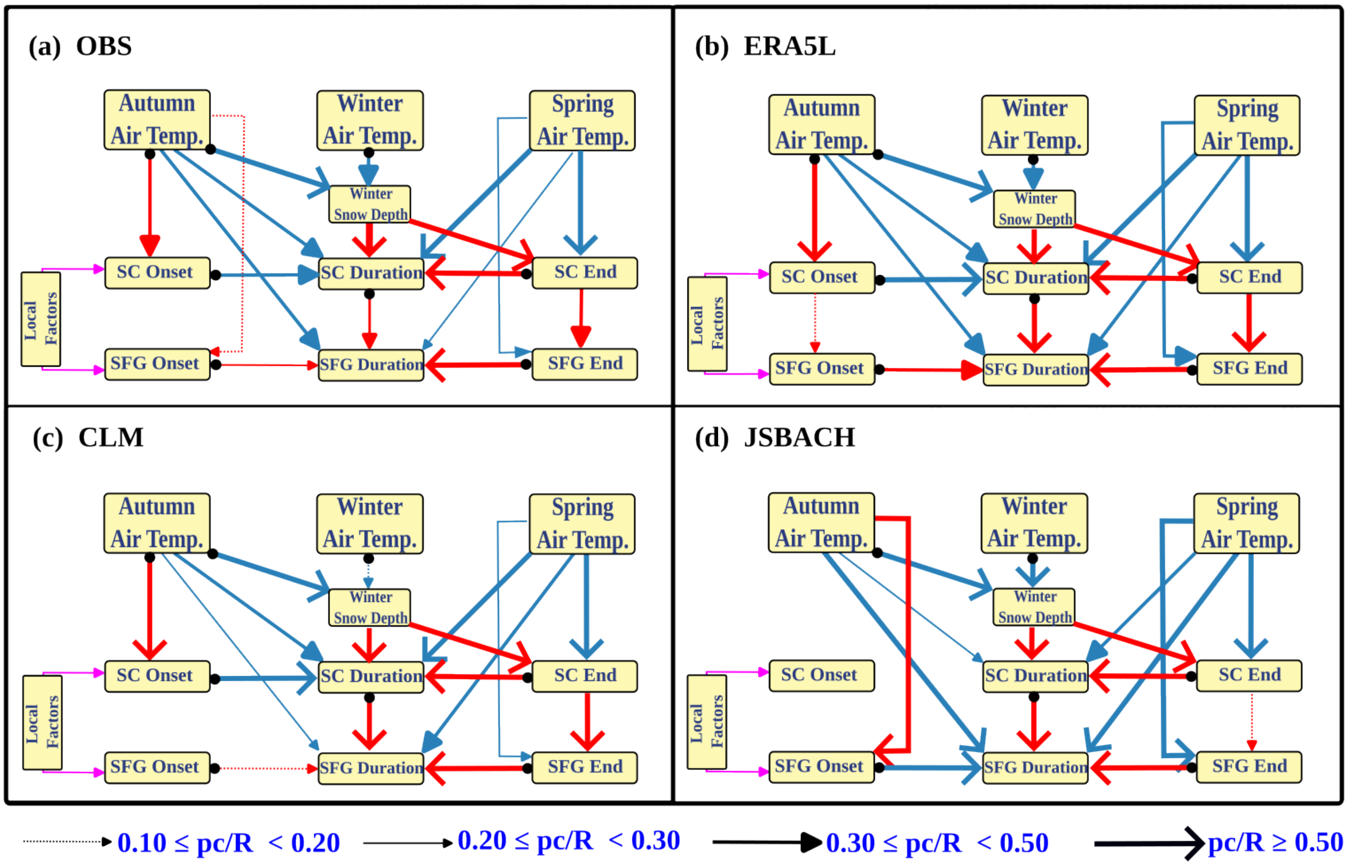


Figure 1: Schematic representation of the statistically significant pearson and partial correlations derived from Appendix Table 3, illustrating the interactions among seasonal air temperature, snow cover, snow depth, and seasonally frozen ground characteristics for (a) OBS, (b) ERA5L, (c) CLM, and (d) JSBACH. Red and blue lines denote positive and negative correlations, respectively, while the magenta-coloured line indicates possible local influences affecting the variables. Filled circles at the arrow origin indicate pearson correlation coefficients; all other values represent partial correlations.

Based on the information we have, we would like to correct our Sect 3.3.4 as follows in the revised manuscript:

Figure 10 shows statistically significant correlation matrices between SFG and snow cover characteristics, seasonal ATs, winter SD, as well as thermal indices for observations and models. To further distinguish direct relationships from shared seasonal influences, partial correlation (pc) analysis was additionally calculated (using `pingouin` package; `https://pingouin-stats.org/generated/pingouin.partial_corr.html`). The resulting partial correlation values are presented in Appendix Table 3, while the corresponding controlling factors used in each analysis are listed in Appendix Table 2. This approach removes the effect of co-varying seasonal controls (autumn AT, spring AT, winter AT and winter SD), allowing a clearer assessment of whether snow cover exerts an independent influence on soil freezing dynamics or whether observed linkages arise primarily through common atmospheric forcing. Figure 1 presents a schematic diagram based on Appendix Table 3, illustrating both pearson correlations (R) and partial correlations (pc) after accounting for shared drivers, thereby highlighting the relative strength and direction of the key interactions among variables.

In observations, SFG onset is not directly controlled by large-scale seasonal conditions. SC_Onset and autumn AT show a direct moderate positive relationship ($R = 0.30$), indicating that warmer autumn temperatures delay snow cover onset, consistent with a temperature-driven control, yet autumn AT exerts only limited direct influence on SFG onset (pc = 0.14), and the mutual partial correlation between SC_Onset and SFG_Onset is negligible (pc = 0.02 after controlling for autumn AT). SC duration is primarily controlled by winter SD (pc = −0.52), spring AT (pc = −0.50), and autumn AT (pc = −0.33). SC_End shows strong dependence on spring AT (pc = −0.66) and winter SD (pc = +0.52), and a moderate relationship with SFG_End after controlling for spring temperature (pc = +0.37). In contrast, SFG_End is correlated with the thawing index (R = -0.50) but is only weakly directly influenced by spring AT once the effect of SC_End is removed (pc = −0.20), highlighting that soil thaw timing is largely mediated through snow cover dynamics rather than temperature alone. SFG duration shows the strongest association with autumn AT (R = −0.52; pc = −0.37 after controlling for SC duration and spring AT), with spring AT retaining a moderate independent effect on SFG duration after controlling for SC duration (pc = −0.21). Altogether, the observations indicate that snow cover governs thaw timing and, together with the intensity of seasonal air temperature, determines the duration of seasonal ground freezing.

ERA5L and CLM broadly reproduce several of the observed dependencies but often amplify the strength of correlations, suggesting an excessive sensitivity of the simulated soil thermal regime to AT and snow controls. Moreover, both models do not account for the high heterogeneity in local conditions present at observation sites, as land surface properties are spatially smoothed, which likely contributes to the strong dependence of snow cover onset on seasonal variables, particularly autumn AT. In addition, precipitation phase partitioning in the models is governed by grid-scale temperature thresholds, making the onset of snow accumulation primarily temperature-driven. Consequently, local factors that modify the near-surface temperature regime, such as vegetation dynamics, soil thermal properties, microtopography, and surface energy exchanges, cannot be represented as specifically as in station observations. The SC_End-SFG_End coupling is particularly inflated ($R > 0.9$; pc = 0.85 in ERA5L and 0.84 in CLM after controlling for spring AT), far exceeding observations. In ERA5L, spring AT retains a stronger direct effect on SFG_End than observed (pc = −0.42), and the SFG duration-SC duration coupling is elevated (R = 0.53 vs. 0.38 in observations), with comparable sensitivity of SFG duration to both autumn AT (pc = −0.32) and spring AT (pc = −0.36). In CLM, SC_Onset is highly sensitive to autumn AT (R = 0.73), and the dominant role of snow accumulation in regulating SC duration is reflected in strong partial correlations with winter SD (pc = +0.69) and spring AT (pc = −0.52). CLM weakens the direct influence of autumn AT on soil freezing duration (pc = −0.20 compared to −0.37 in observations), while slightly strengthening the independent role of spring AT (pc = −0.33 compared to −0.21 in observations). Collectively, the amplified correlations indicate that excessive snow accumulation in ERA5L and CLM produces a tighter-than-observed coupling between snowpack evolution and subsurface freezing, thereby reducing the influence of other environmental controls on SFG variability.

JSBACH on the other hand, exhibits a distinctly different relationship structure, likely reflecting its substantial snow depth underestimation. SC_Onset shows no dependence on autumn AT (R = 0.07), while SFG_Onset is strongly tied to autumn AT (pc = 0.71) with no meaningful SC_Onset-SFG_Onset relationship (pc = -0.04). At the thaw end, the SC_End-SFG_End partial correlation weakens (pc = 0.18 vs. 0.37 in observations), while spring AT's direct effect on SFG_End remains very strong (pc = -0.68), indicating that spring warming forces soil thaw without mediation through snow cover. Consequently, both autumn AT (pc = -0.52) and spring AT (pc = -0.54) exert substantially stronger independent control on SFG duration than observed, bypassing the buffering role of snow insulation. A notable sign reversal in JSBACH where SFG_Onset shows a strong negative relationship with SFG duration (R = -0.77), alongside a strong positive SFG_End-SFG duration relationship (R = +0.84), indicates that both earlier onset and delayed end contribute to longer frozen ground duration, whereas in observations and the other models, SFG duration is primarily controlled through the end date.

The model's limited snow accumulation and relatively shallow snowpack fail to adequately decouple the soil from atmospheric temperature variations, resulting in a more direct soil thermal response to AT. Together, these results highlight that observed SFG characteristics arise from a balanced interaction between snow dynamics and seasonal thermal/freeze energy, whereas models diverge in how they distribute control between AT and snow insulation. ERA5L and CLM exaggerate the thermal decoupling effect of snow, while JSBACH underrepresents it, leading to contrasting biases in the simulated freeze-thaw behavior.

Minor comments

Comment 1: Lines 91-92: What quantities are aggregated to the grid scale (e.g., energy, hydrology, carbon)?

Response: As stated in Lines 91-92, JSBACH uses a tiling approach in which different land-cover types are simulated separately and then aggregated to the grid scale. Additive quantities, such as flux densities (energy, water, carbon) and albedo, are aggregated using area-weighted averaging based on tile fractions. However, non-additive quantities, such as surface temperature and roughness length, are treated differently: surface temperature is computed at the grid-box level to ensure energy conservation, while subsurface heat and moisture fluxes are calculated at the tile level and averaged at the end of each time step. We will update Lines 91-92 to clarify this.

Comment 2: Line 146: Why was JSBACH run without the two-phase spin-up?

Response: We thank the reviewer for this question. The spin-up procedures for both models are described in detail in our response to Reviewer #1 (Comment 1), where we clarify the differences in spin-up strategies between CLM and JSBACH. We will also add a brief paragraph in the revised manuscript to improve clarity.

Comment 3: Line 112: Please specify the version of CLM used.?

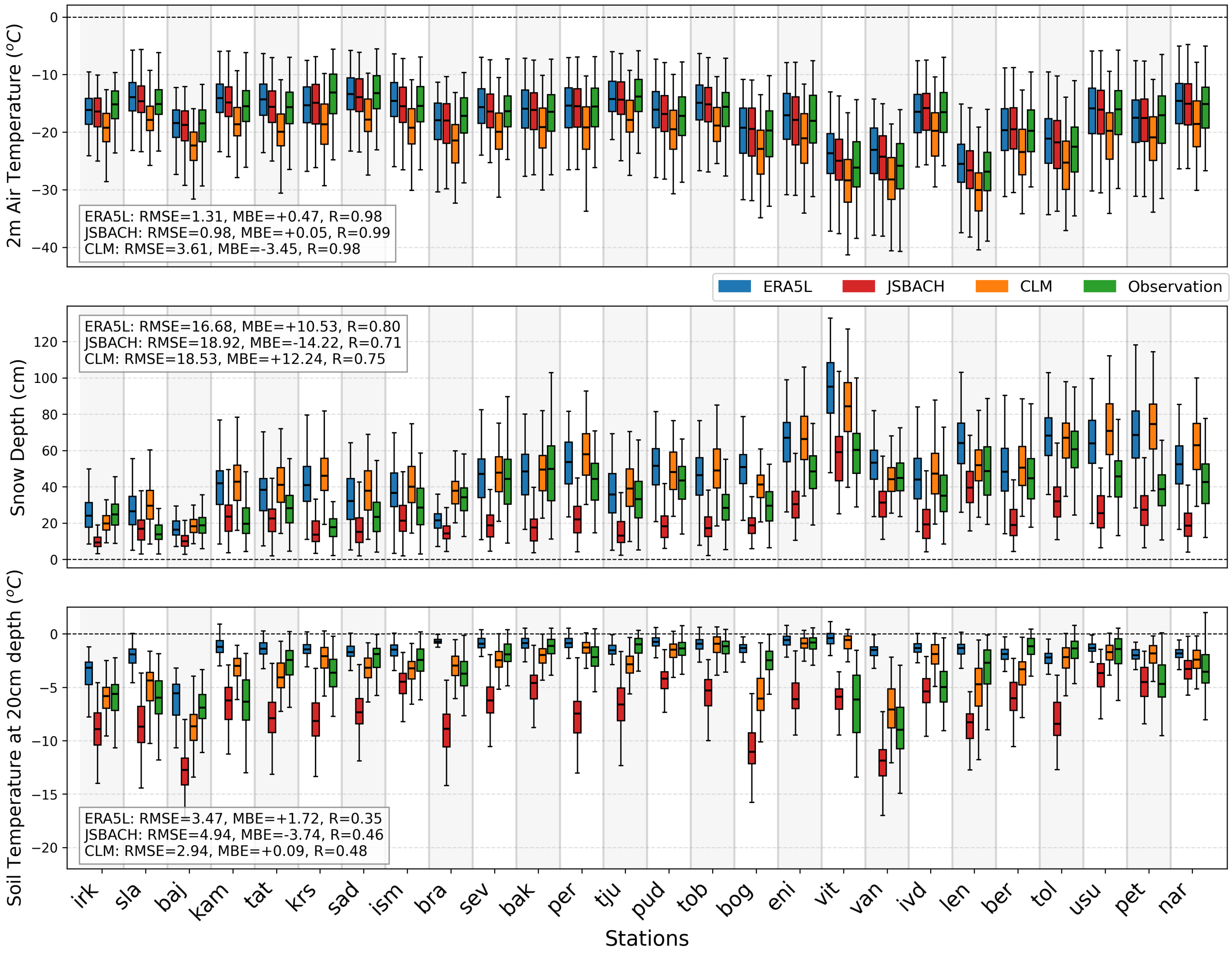
Response: The version of CLM used in this study has been specified in the manuscript (Line 120: CTSM5.1.dev128_cm3_v1 (`https://github.com/CMCC-Foundation/CTSM`; hereafter referred to as CLM)).

Comment 4: Section 3.1: Is this section better suited to the Methods rather than Results??

Response: We thank the reviewer for this helpful suggestion. We agree that this section is better suited to the Methods. Following this and similar suggestions from other reviewers, we will move Section 3.1 to the end of the methodology section in the revised manuscript.

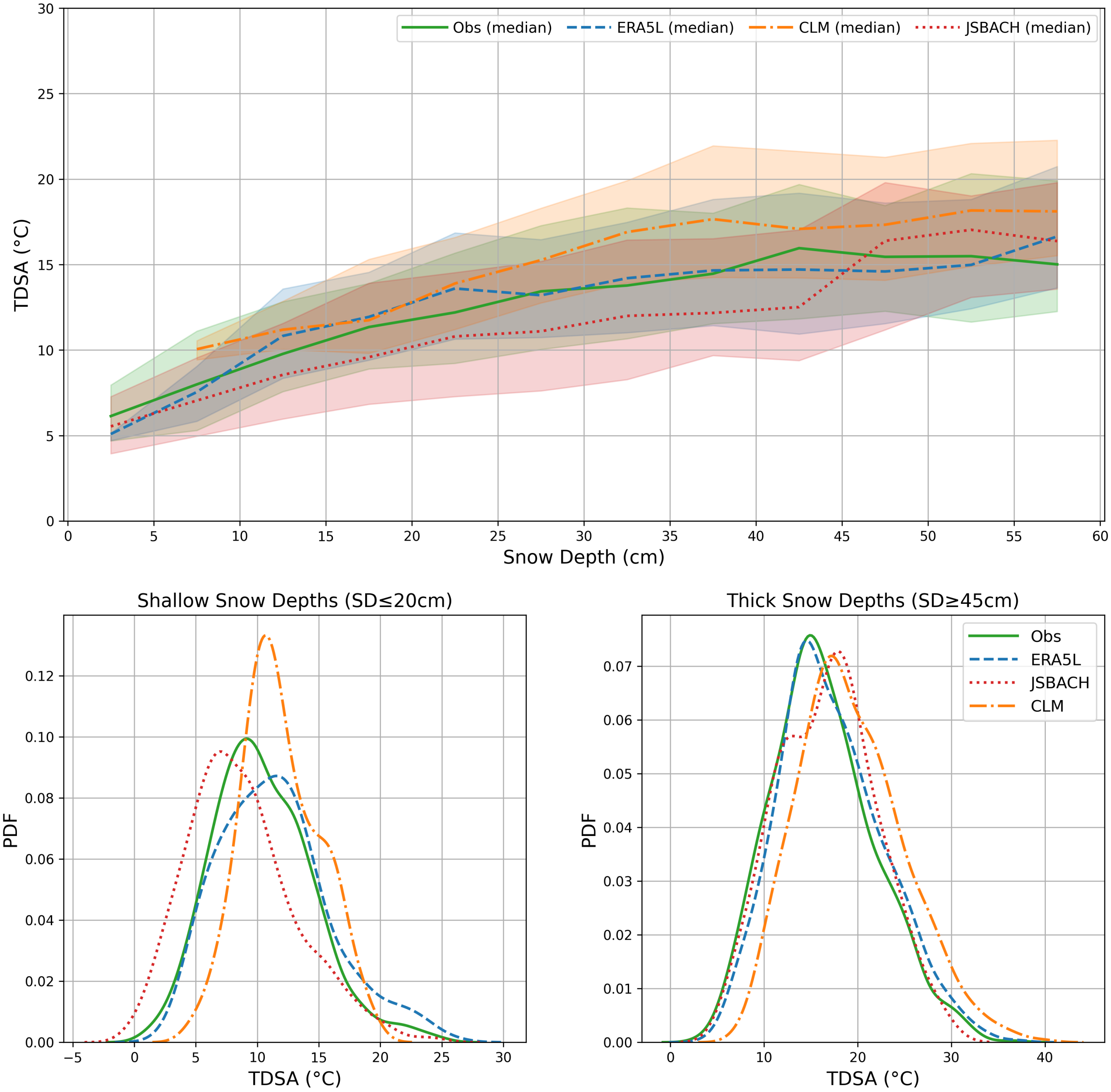
Comment 5:Figure 6: Showing results for all individual sites may reduce readability. Consider presenting site-averaged results or aggregating sites based on a classification scheme.

Response: We appreciate the concern regarding readability and acknowledge that presenting all individual sites in a single figure can be visually dense. However, we have chosen to retain this figure to allow the reader to directly assess site-specific variations and to better understand how underestimation or overestimation of snow depth influences model performance at individual locations. We believe this individual-site view provides valuable insight that would be lost through aggregation. In addition, we have provided aggregated performance metrics (RMSE, MBE, and correlation) across all sites, which offer a concise summary of overall model behaviour. To further improve readability, we plan to revise the figure layout by introducing clearer separation between sites (e.g., distinct vertical groupings), as illustrated in the figure below, which will be included in the revised manuscript.



Comment 6:Figure 7: It may improve clarity to show observations as solid black lines and model results as colored dashed lines.

Response: Following the reviewer's recommendation, we updated the figure by keeping the observational line as solid black and converting all model lines to different dashed styles while retaining the original color scheme. During this revision, we found that the shaded representation of snow-related effects reduced the visual clarity of overlapping curves, particularly against the solid black observational line. In addition, the use of different line styles improves accessibility and makes the distinction between datasets clearer, particularly for colorblind readers. This revision has been implemented in the updated figure shown below and will be included in the revised manuscript.



References

Koven, C. D., Riley, W. J., and Stern, A.: Analysis of Permafrost Thermal Dynamics and Response to Climate Change in the CMIP5 Earth System Models, Journal of Climate, 26, 1877-1900, https://doi.org/10.1175/JCLI-D-12-00228.1, 2013.

Slater, A. G. and Lawrence, D. M.: Diagnosing Present and Future Permafrost from Climate Models, Journal of Climate, 26, 5608-5623, https://doi.org/10.1175/JCLI-D-12-00341.1, 2013.

González-Rouco, J. F., Steinert, N. J., García-Bustamante, E., Hagemann, S., De Vrese, P., Jungclaus, J. H., Lorenz, S. J., Melo-Aguilar, C., García-Pereira, F., and Navarro, J.: Increasing the Depth of a Land Surface Model. Part I: Impacts on the Subsurface Thermal Regime and Energy Storage, Journal of Hydrometeorology, 22, 3211-3230, https://doi.org/10.1175/JHM-D-21-0024.1, 2021.