Response to Reviewer#1 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:

To me, the spin up configuration has flaws. The initial condition of soil, including the thermal and hydrological state of soil, especially for the seasonally frozen soil with water phase change in it, are critically important. According to the manuscript, the authors use a historical simulation from 1850 to 1940 forced by GSWP3, which is a different forcing dataset from ERA5. The spin up is basically a transient simulation, so that to me there is no way to judge if the spin up is completed. Authors should provide additional information on whether the thermal state of soil and atmosphere has reached the equilibrium by the end of the spin up, otherwise the simulated soil temperature, as well as the snow accumulation manners, are not comparable to the observation/ERA-land. Furthermore, the author mentioned that the JSBACH simulation is run without the two-phase spin up like the CLM5 simulation did. I am wondering how the initial conditions used by JSBACH and CLM5 differ from each other. The difference in initial condition could make the soil state totally different. Additionally, to my knowledge the GSWP3 data covers the period of 1901-2014 (`https://svn-ccsm-inputdata.cgd.ucar.edu/trunk/inputdata/atm/datm7/atm_forcing.datm7.GSWP3.0.5d.v1.c170516/TPHWL/`). So how exactly do authors use this for the spin up from 1950 to 1899?

Response:

We sincerely thank the reviewer for this careful and important comment regarding the spin up strategy and initialization procedure. We fully acknowledge that the description in the original manuscript was not sufficiently detailed and may have created confusion regarding the model equilibration procedure. We will substantially revise the manuscript to clarify these aspects based on the following information

- **CLM and JSBACH spin up method:**

  We agree with the reviewer that demonstrating thermal and hydrological equilibrium at the end of spin up is particularly important for seasonally frozen soils, where freeze-thaw processes are sensitive to initial land states. We clarify that CLM5.1 was initialized following the standard two-phase spin up procedure described in Lawrence et al. (2019); Koven et al. (2013), and implemented in the CMCC seasonal prediction system (Gualdi et al., 2020). In the first phase, the model was brought to equilibrium representative of year-1850 conditions using the standard CLM spin up protocol. This phase consisted of 400 years with accelerated decomposition to equilibrate slow carbon and nitrogen pools, followed by an additional 800 years in normal mode. The second phase was a historical run from 1850, which was achieved by integrating over a repeating 20-year period of the GSWP3 forcing dataset (1901-1920) (Lawrence et al., 2019), while prescribing fixed preindustrial boundary conditions, including constant atmospheric CO_2, nitrogen deposition, aerosol deposition, land use, and no wood harvest. This is the standard CLM approach for generating a

stable preindustrial land state when meteorological forcing prior to 1901 is unavailable. From 1901 onward, the model was forced with true time-varying GSWP3 data, which in our setup covers the period 1901-1940. To ensure a smooth transition from GSWP3 to ERA5 forcing, we additionally performed 10 cyclical runs using climatological ERA5 forcing for 1941-1950, which reduced potential drift caused by the forcing dataset change.
In contrast, JSBACH was run continuously from 1941 to 2022 without an equivalent multi phase spin up, initialized from JSBACH3 initial-condition files remapped from the Gaussian grid to the ICON R02B06 grid, and forced with ERA5 reanalysis data at a 3-hourly temporal resolution. Previous JSBACH site-level study has considered a spin up period of 30 years sufficient to reach thermal and hydrological equilibrium in frozen ground simulations (Ekici et al., 2014). Since our analysis period begins in 1986, JSBACH had been integrated for 45 years (1941-1985) prior to evaluation, exceeding this threshold and ensuring that the influence of the initial conditions on the analysed variables is negligible. As dynamic nitrogen cycling is not included in either CLM or JSBACH, slowly evolving nitrogen pools do not impose additional spin up requirements.

- **Evidence of equilibrium - supplementary figures:**

  To directly address the concern regarding whether equilibrium was reached, we will add the following figures in the supplementary material showing the temporal evolution of annual mean soil temperature and soil moisture at the deepest soil layer, and annual mean snow depth over a selected region (Figure 1) for the full simulation period 1941-2022 (shown in Figure 2). Figure 2 demonstrate that both CLM and JSBACH achieve a stable equilibrium state well before the analysis period begins in 1986.

  As shown in Figure 2, CLM5 starts the transient simulation in an already stable state, reflecting the effectiveness of the two-phase spin up procedure. For JSBACH, which was initialized directly in 1941 without a comparable multi-phase spin up, an initial adjustment period of approximately 10-15 years is visible at the beginning of the simulation. After this period, the simulated soil temperature and snow depth primarily show interannual variability and longer-term trends related to climate warming.

  We agree that the initial conditions of the two models were different owing to their distinct model structures and initialization frameworks. However, since the analysis period starts in 1986, both models had already been run for several decades before evaluation (45 years from 1941 to 1985). Therefore, the direct influence of the initial conditions on the analysed frozen-ground variables is expected to be small during the evaluation period, as soil thermal and hydrological memory typically dissipates over such timescales (Yang and Zhang, 2016; Koster and Suarez, 2001; Vinnikov et al., 1996). But since it is important for the readers to know it so we are planning to acknowledge that “Differences in spatial resolution, grid-cell heterogeneity, initial state and boundary condition datasets describing soil properties likely contribute to some of the simulated biases.” And we will clarify this point in the revised manuscript (Line 323).

Lines (141-147) will be updated in the revised manuscript to give more clarity on the spin up procedure as follows:
For CLM, the BGC-CROP biogeochemistry crop model was activated (not used for JSBACH), and the model was initialized using a standard two-phase spin up procedure (Koven et al., 2013; Lawrence et al., 2019). First, a preindustrial spin up was performed with 400 years of accelerated decomposition

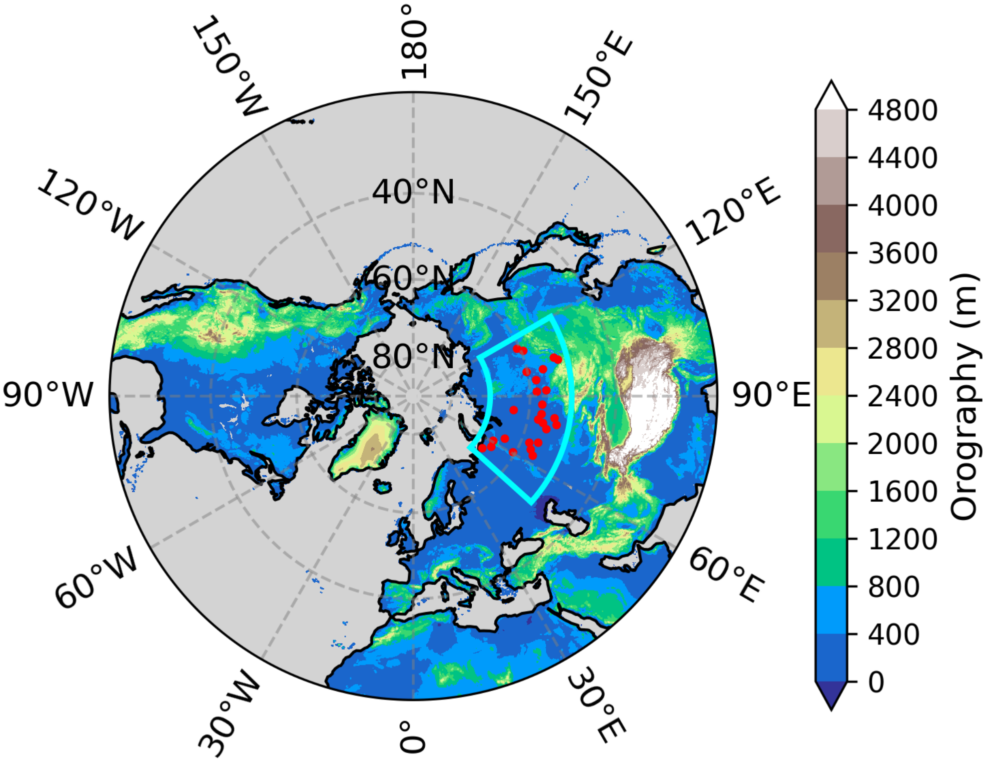


Figure 1: Selected study region for the equilibrium analysis shown by the cyan color box on the orography map.

followed by 800 years in normal mode to equilibrate slow carbon and nitrogen pools. Equilibrium conditions for the year 1850 were obtained by integrating over a repeating 20-year cycle (1901-1920) of Global Soil Wetness Project (GSWP3) meteorological forcing under fixed preindustrial boundary conditions. From 1901 onwards, the model was forced with GSWP3 data until 1940, followed by 10 cyclical runs using climatological ERA5 forcing for 1941-1950 to ensure a smooth transition between forcing datasets, with the final spin up state serving as the initial condition for the transient simulation from 1941 to 2022 using 3-hourly ERA5 data. In contrast, JSBACH was run continuously from 1941 to 2022 without an equivalent two-phase spin up. The analysis for both models focused on 1986-2022 to match the availability of observational datasets. Supplementary Figures S1-S2 show the temporal evolution of deep soil temperature, soil moisture and snow depth, confirming that both models reached stable states before the analysis period.

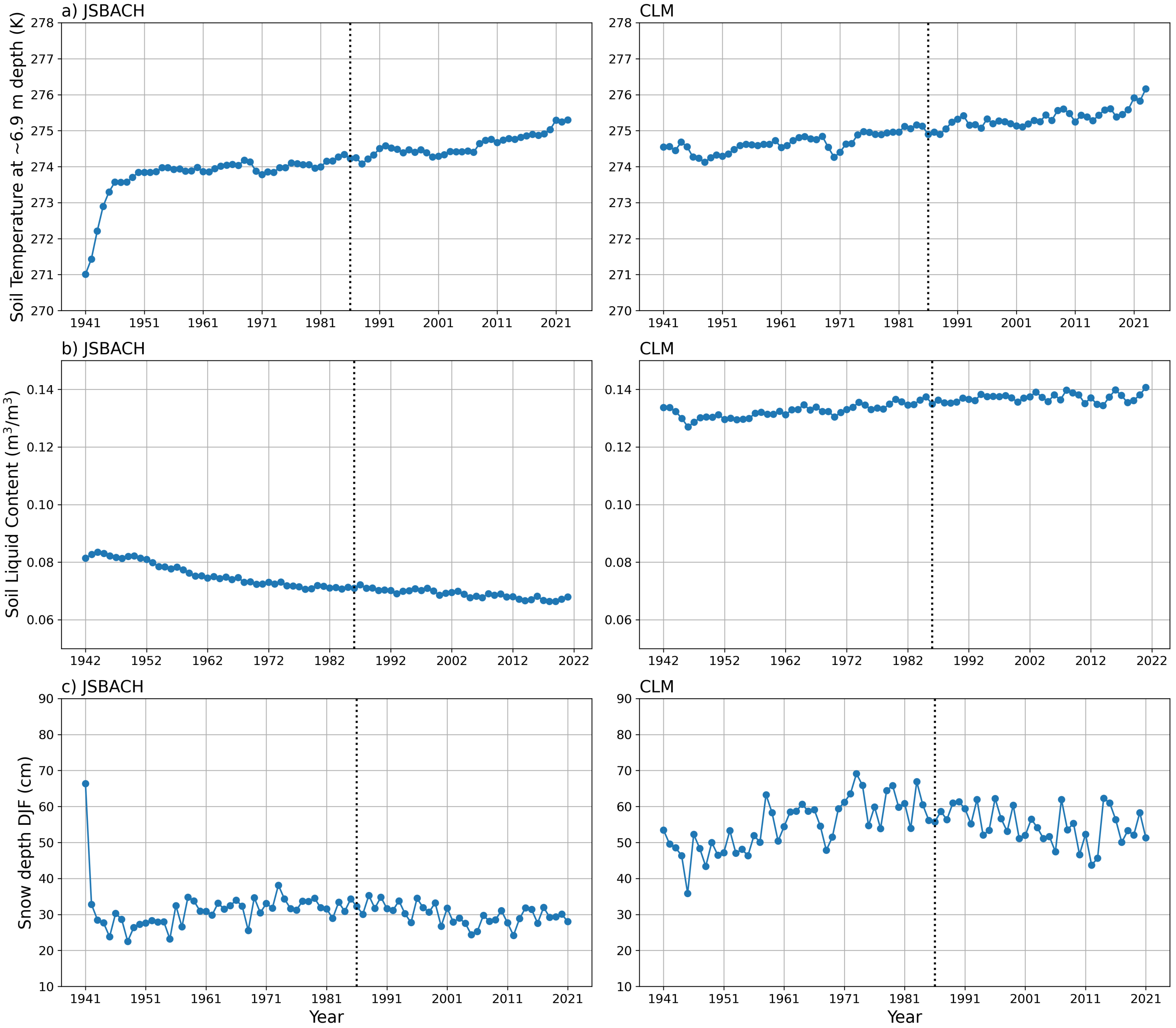


Figure 2: Time series of annual mean (a) soil temperature at 6.9 m depth, (b) volumetric soil liquid content for the whole soil column (m^3/m^3), and (c) snow depth over the regions shown in Figure S1, simulated by JSBACH (first column) and CLM5 (second column) for the period 1941-2022. The dashed vertical line indicates the start of the analysis period (1986).

Comment 2:

According to Table 1, the JSBACH model only has 5 soil layers. To me it is a little bit too few to present the soil temperature profile in permafrost area. Authors should consider add the information of soil layer structure (the depth and thickness of each soil layer) of the two models in the manuscript. Also, the CLM model offers a detailed soil layer structure tuned for permafrost modeling (49 soil layers for 0-9.85 m of soil and 5 bedrock layers). By the way, the definition of permafrost is deviated from that by the International Permafrost Association (the soil with a temperature colder than 0 °C (32 °F) continuously for two or more years). Authors should explain and verify the validity of this definition and how this definition leads to frozen ground extent in the JSBACH model.

Response:

We thank the reviewer for these two distinct and important points. We address them separately below.

We agree that JSBACH uses a relatively coarse vertical discretization (5 soil layers), especially when compared to CLM, which includes a much finer vertical structure (20 soil layers in the soil column extending to 8.6 m depth, plus 5 bedrock layers until about 50 m). Both models were chosen because they are part of Earth System Models (ESMs) that are used for climate projections and climate predictions. As such, their ability to realistically represent seasonally frozen soils is of considerable interest. This comparison provides insight into their respective strengths and limitations, particularly regarding whether a coarser soil vertical discretization can reproduce similar large scale frozen ground patterns as a higher resolution model. Importantly, the aim of this study is not to benchmark one model against the other in terms of structural complexity, but rather to evaluate how two models with fundamentally different configurations, one highly detailed and widely used in frozen ground research and one with a comparatively simpler structure, represent seasonally frozen ground characteristics. Including both models in this study is therefore useful, as it allows us to assess how differences in model complexity and vertical discretization influence the simulation of frozen ground. We will 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))

Regarding the permafrost definition, we agree that our definition of Permafrost (PEFT) differs from the standard International Permafrost Association (IPA) definition, which requires ground temperatures to remain below 0°C continuously for two or more consecutive years. We clarify that our classification defines a grid cell as PEFT if, during the 1986-2022 analysis period (37 years), at least one soil layer remains frozen year-round for more than 50% of the years. In practice, this means that a cell must exhibit year-round frozen conditions in at least 19 out of the 37 years. The requirement of $\geq$19 years with year-round frozen ground ensures that at least one consecutive year of complete freezing occurs (thereby satisfying the baseline condition of the IPA definition), while also capturing the full range of permafrost persistence up to all 37 years being frozen. Therefore, A threshold of >50% guarantees that freeze conditions are climatologically persistent rather than occasional or episodic, so a necessary condition when working with multi-decadal model outputs where interannual variability is high.

Apart from the two-consecutive-year criterion, a similar approach as ours using individual simulation year criteria has been used by Langer et al. (2024) based on a thaw depth parameter, and previous large-scale permafrost assessments have also used alternative diagnostics based on active layer thickness (Dankers et al., 2011), thermal index (Slater and Lawrence, 2013), or threshold-based probabilistic occurrence of permafrost (Gruber, 2012; Burke et al., 2020).

We will revise the manuscript to explicitly clarify the PEFT definition and justify our methodological choice as follows (at Line 246):

Here, PEFT is defined as grid cells where at least one soil layer remains frozen year-round in >50% of the 1986-2022 simulation years (i.e., $\geq$19 of 37 years), representing persistent frozen-ground conditions in a multi-decadal framework. Although the International Permafrost Association (IPA) conventionally defines PEFT as ground remaining below 0 ^{o}C for two or more consecutive years (Brown et al., 1997), alternative diagnostics are commonly used in gridded model assessments (Steinert et al., 2024; Langer et al., 2024; Gruber, 2012; Slater and Lawrence, 2013; Burke et al., 2020). SFG denotes areas experiencing seasonal freezing for at least 15 days per year.

Comment 3:

Regarding snow depth, it should be noted that it is closely related to how the model deals with solid precipitation. The GSWP3 forcing only has total precipitation (rain+snow), and it is the land surface model that deals with the precipitation partitioning. In this way, other than the snow accumulation modeling, authors should also elaborate on the difference, if any, of precipitation partitioning between the two models in the methodology section.

Response:

We thank the reviewer for pointing this out.

Both JSBACH and CLM5 uses temperature-based linear partitioning methods, but with different threshold temperature limits, as shown below, which we plan to add as an additional row in Table 1. Therefore, we agree that differences in simulated snowfall can arise not only from snowpack physics, but also from the precipitation partitioning schemes used by the two models.

Model features	**JSBACH**	**CLM**
Precipitation (P) partitioning using air temperature (T ^{o}C)	$\text{snow} = \begin{cases} P & T < -1.1 \\ P\dfrac{3.3 - T}{4.4} & -1.1 \leq T \leq 3.3 \\ 0 & T > 3.3 \end{cases}$ rain = P − snow (Wigmosta et al., 1994)	$\text{snow} = \begin{cases} P & T < 0 \\ P\dfrac{2 - T}{2} & 0 \leq T \leq 2 \\ 0 & T > 2 \end{cases}$ rain = P − snow (Lawrence et al., 2019)

Such differences are generally most evident during transition seasons, particularly at the beginning and end of winter, when near-surface air temperature lies within the mixed-phase range between the minimum and maximum threshold temperatures used for rain-snow partitioning. Under these conditions, model specific threshold values can lead to some differences in snowfall fraction, which should be acknowledged in the main text. We have also evaluated the rainfall and snowfall fractions in both models. The plot below shows only small differences in monthly climatological accumulated rainfall and snow fractions over the selected study region (Figure 1) during 1986-2022, indicating that the overall partitioning differences are minor. We will include this supporting figure in the revised material, if needed. We will also add a brief note at Line 229 of the revised manuscript stating:

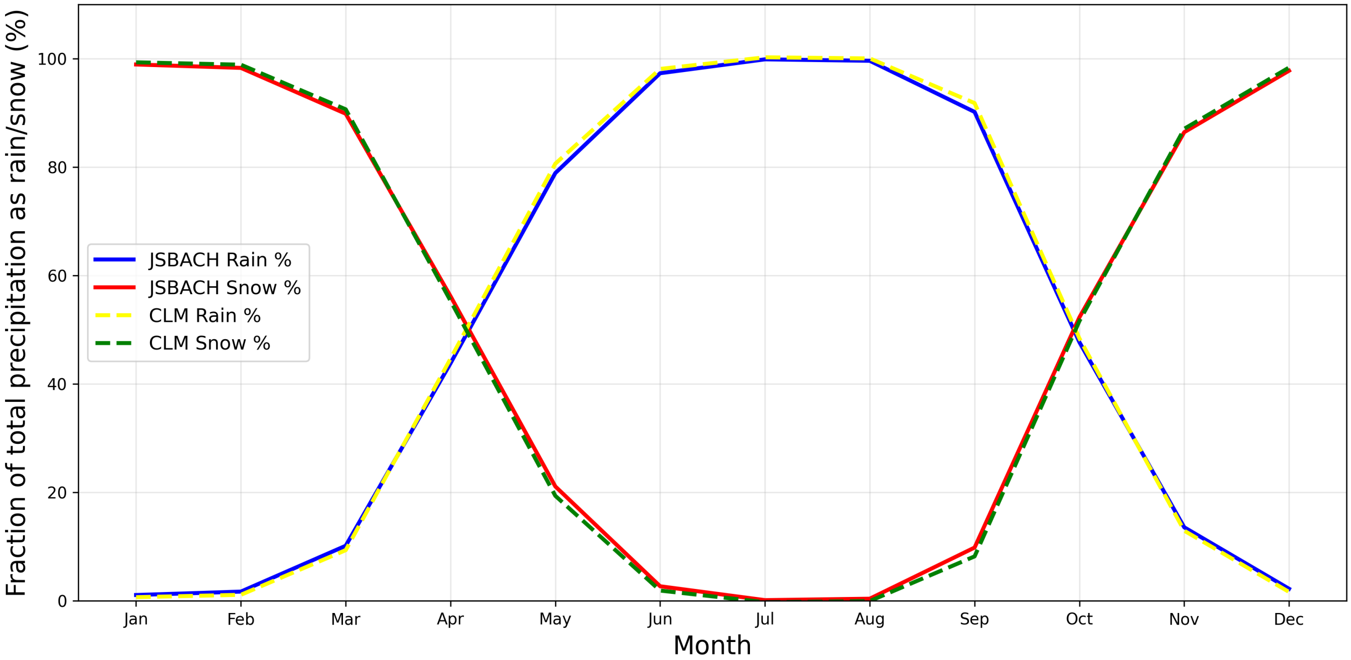


Figure 3: Monthly climatological accumulated rainfall and snowfall fractions averaged over the selected region (Figure 1).

“Since JSBACH and CLM models use different precipitation partitioning schemes (Table 1), small differences in simulated snow related variables are expected, particularly during transitional seasons when near-surface air temperatures fall within the model specific rain-snow threshold range.”

Minor comments

Comment 1: Section 3.1 This part should be moved to the methodology section.

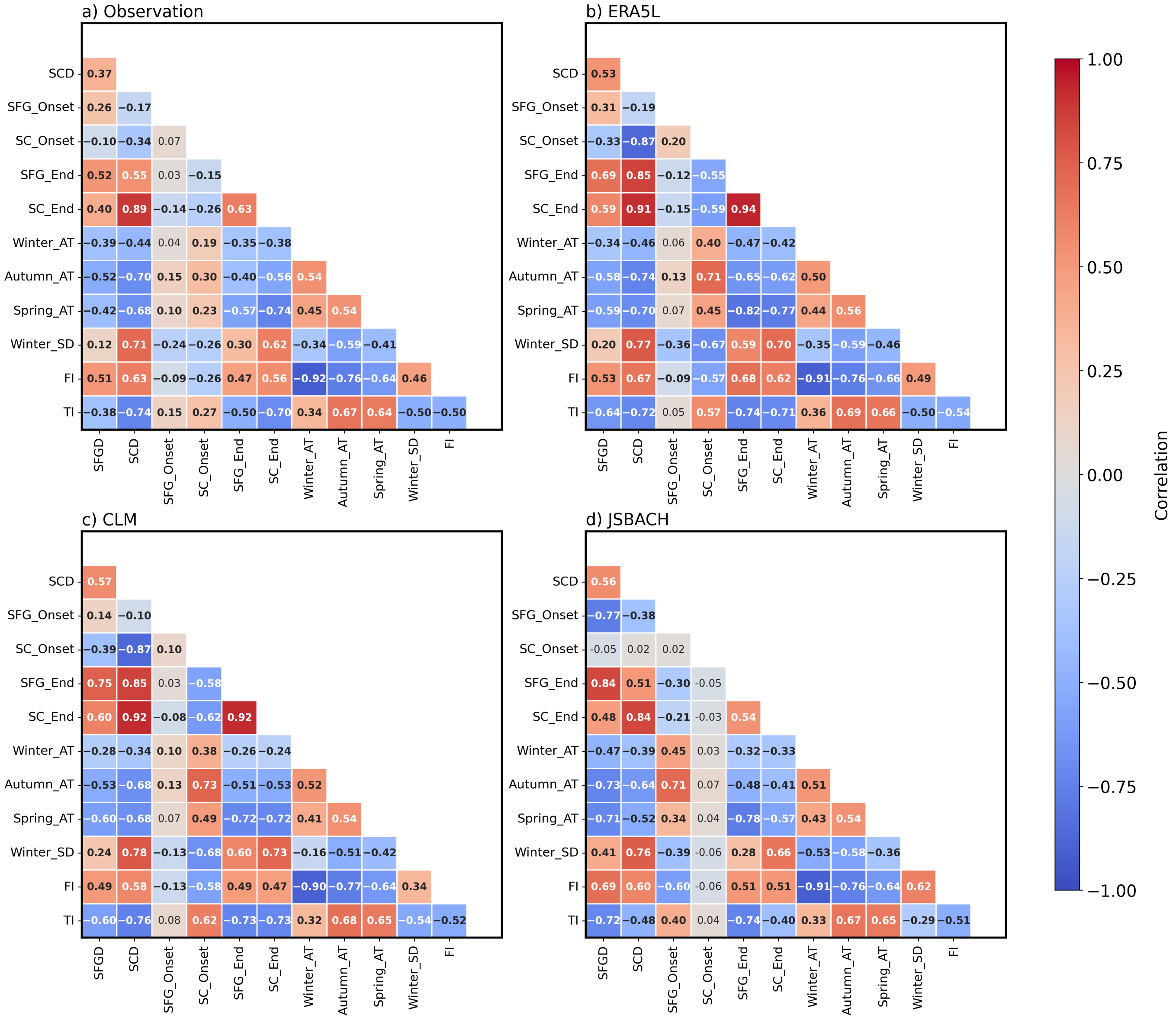
Response:

Following your and the other reviewers’ suggestions, we will move this section to the end of the methodology section.

Comment 2: Figure 10: I suggest adding significance test on top of these linear correlation coefficients.

Response:

Thank you for your suggestion. The following significance test plot will replace the previous one in the revised manuscript, with statistically significant values highlighted in bold.



References

Koven, C. D., Riley, W. J., Subin, Z. M., Tang, J. Y., Torn, M. S., Collins, W. D., Bonan, G. B., Lawrence, D. M., and Swenson, S. C.: The effect of vertically resolved soil biogeochemistry and alternate soil C and N models on C dynamics of CLM4, Biogeosciences, 10, 7109-7131, `https://doi.org/10.5194/bg-10-7109-2013`, 2013.

Brown, J., Ferrians, O. J. J., Heginbottom, J. A., and Melnikov, E. S.: Circum-Arctic map of permafrost and ground ice conditions, US Geological Survey, Reston, VA, 1997.

Lawrence, D. M., Fisher, R. A., Koven, C. D., Oleson, K. W., Swenson, S. C., Bonan, G., Collier, N., Ghimire, B., Van Kampenhout, L., Kennedy, D., Kluzek, E., Lawrence, P. J., Li, F., Li, H., Lombardozzi, D., Riley, W. J., Sacks, W. J., Shi, M., Vertenstein, M., Wieder, W. R., Xu, C., Ali, A. A., Badger, A. M., Bisht, G., Van Den Broeke, M., Brunke, M. A., Burns, S. P., Buzan, J., Clark, M., Craig, A., Dahlin, K., Drewniak, B., Fisher, J. B., Flanner, M., Fox, A. M., Gentine, P., Hoffman, F., Keppel-Aleks, G., Knox, R., Kumar, S., Lenaerts, J., Leung, L. R., Lipscomb, W. H., Lu, Y., Pandey, A., Pelletier, J. D., Perket, J., Randerson, J. T., Ricciuto, D. M., Sanderson, B. M., Slater, A., Subin, Z. M., Tang, J., Thomas, R. Q., Val Martin, M., and Zeng, X.: The Community Land Model Version 5: Description of New Features, Benchmarking, and Impact of Forcing Uncertainty, Journal of Advances in Modeling Earth Systems, 11, 4245-4287, `https://doi.org/10.1029/2018MS001583`, 2019.

Zhang, T., Heginbottom, J. A., Barry, R. G., and Brown, J.: Further statistics on the distribution of permafrost and ground ice in the Northern Hemisphere1, Polar Geography, 24, 126-131, https://doi.org/10.1080/10889370009377692, 2000.

Gualdi, S., Borrelli, A., Cantelli, A., Davoli, G., del Mar Chavesmontero, M., Masina, S., Navarra, A., Sanna, A., and Tibaldi, S.: The new CMCC operational seasonal prediction system, CMCC Technical Notes, TN0288, `https://www.cmcc.it/wp-content/uploads/2020/09/TN0288-csp-09-2020-1.pdf`, 2020.

Steinert, N. J., Debolskiy, M. V., Burke, E. J., García-Pereira, F., and Lee, H.: Evaluating permafrost definitions for global permafrost area estimates in CMIP6 climate models, Environ. Res. Lett., 19, 014033, https://doi.org/10.1088/1748-9326/ad10d7, 2024.

Koster, R. D. and Suarez, M. J.: Soil Moisture Memory in Climate Models, J. Hydrometeor, 2, 558-570, https://doi.org/10.1175/1525-7541(2001)002%3C0558:SMMICM%3E2.0.CO;2, 2001.

Vinnikov, K. Y., Robock, A., Speranskaya, N. A., and Schlosser, C. A.: Scales of temporal and spatial variability of midlatitude soil moisture, J. Geophys. Res., 101, 7163-7174, https://doi.org/10.1029/95JD02753, 1996.

Ekici, A., Beer, C., Hagemann, S., Boike, J., Langer, M., and Hauck, C.: Simulating high-latitude permafrost regions by the JSBACH terrestrial ecosystem model, Geoscientific Model Development, 7, 631-647, https://doi.org/10.5194/gmd-7-631-2014, 2014.

Yang, K. and Zhang, J.: Spatiotemporal characteristics of soil temperature memory in China from observation, Theor Appl Climatol, 126, 739-749, https://doi.org/10.1007/s00704-015-1613-9, 2016.

Burke, E. J., Zhang, Y., and Krinner, G.: Evaluating permafrost physics in the Coupled Model Intercomparison Project 6 (CMIP6) models and their sensitivity to climate change, The Cryosphere, 14, 3155-3174, https://doi.org/10.5194/tc-14-3155-2020, 2020.

Dankers, R., Burke, E. J., and Price, J.: Simulation of permafrost and seasonal thaw depth in the JULES land surface scheme, The Cryosphere, 5, 773-790, https://doi.org/10.5194/tc-5-773-2011, 2011.

Gruber, S.: Derivation and analysis of a high-resolution estimate of global permafrost zonation, The Cryosphere, 6, 221-233, https://doi.org/10.5194/tc-6-221-2012, 2012.

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.

Langer, M., Nitzbon, J., Groenke, B., Assmann, L.-M., Schneider Von Deimling, T., Stuenzi, S. M., and Westermann, S.: The evolution of Arctic permafrost over the last 3 centuries from ensemble simulations with the CryoGridLite permafrost model, The Cryosphere, 18, 363-385, https://doi.org/10.5194/tc-18-363-2024, 2024.

Wigmosta, M. S., Vail, L. W., and Lettenmaier, D. P.: A distributed hydrology-vegetation model for complex terrain, Water Resources Research, 30, 1665-1679, https://doi.org/10.1029/94WR00436, 1994.