## REVIEWER 1

**In this revised manuscript, the authors considered my comments from the first round of revisions. The manuscript has substantially improved. Particularly, the additional analysis on the changes in SWE based on the changes in land cover (LC) is welcome. The authors also found a mistake in the filtering procedure for ESA CCI SWE. With this new filtering procedure, the modelled SWE values are closer to the ESA CCI SWE.**

**However, I find that now the improvements in simulated SWE simulated with the new LC are marginal. Indeed, the authors find that the SWE RMSD is reduced by less than 2 mm using the new LC. This is a result itself, but I think the manuscript should be rephrased to emphasis that the change of LC only slightly changes the overall model performance (similar for LST).**

**RESPONSE 1.1:**

We thank the reviewer for this important observation and agree that, after correcting the ESA CCI filtering, the domain-averaged improvements obtained with the updated land cover are smaller. We have therefore revised the manuscript throughout to avoid overstating the impact of ECOCLIMAP-SG. In particular, the Results section now consistently describes the improvements as modest and mainly associated with reductions in systematic errors rather than improvements in temporal agreement. We also introduced the calculation of the normalised RMSD (nRMSD) for SWE and LST. The nRMSD is the RMSD divided by the mean SWE and LST values above a certain threshold. Threshold values of 15 mm and 5°C are used for SWE and LST, respectively.

In Section 4,2, we replaced:

'On the other hand, RMSD, ubRMSD, and MB are notably reduced in New LC simulations: from 34.7 mm to 32.9 mm, 34.1 mm to 32.6 mm, and 6.4 mm to 4.9 mm, respectively. This reduction in RMSD is caused by the decrease of both MB and ubRMSD values. This suggests that the more detailed and updated land cover dataset better constrains snow accumulation processes in ISBA.'

by

***The RMSD, ubRMSD, nRMSD, and MB are slightly reduced in the New LC simulations, decreasing from 34.7 mm to 32.9 mm, 34.1 mm to 32.6 mm, 62 % to 59 %, and 6.4 mm to 4.9 mm, respectively. While the reduction in RMSD, ubRMSD and MB is less than 2 mm, the reduction in RMSD is 3 %.***

In addition, we also expanded Section 5.1 to explain that the improvements are spatially heterogeneous, with substantial local improvements occurring only for specific land-cover transitions. Consequently, the domain-average statistics improve only slightly although clear local improvements are observed. The new paragraph added to Section 5.1 is presented in the answer to the Major Comment 2 below,

**In addition, I find the SWE RMSD over the domain quite large (values between 33 and 35 mm, Table 1), when the SWE values across the domain are mostly less than 100 mm (Figure S7). I think the overall performance of the model should be better discussed and avenues to improve it in the future suggested. Based on these two comments, I would recommend major revisions of the revised manuscript.**

**RESPONSE 1.2:**

We thank the reviewer for this valuable comment. We agree that, although the updated land cover dataset improves the SWE simulations, the overall SWE RMSD remains relatively large at the domain scale and deserves further discussion.

To address this comment, we have expanded the Discussion section 5.1 by explicitly acknowledging that updating land cover alone cannot remove all sources of SWE uncertainty. We now discuss that snow simulation errors are influenced by uncertainties in atmospheric forcing, snowfall and precipitation phase partitioning, snow physical parameterizations, vegetation structural properties, and topographic effects. We further emphasize that land cover updates should be considered one component of a broader strategy for improving snow simulations and have added a paragraph outlining possible future developments, including improved atmospheric forcing, refined snow physics, higher-resolution vegetation datasets, and satellite-based snow data assimilation.

We added this at the end of Section 5.1 to answer both Major comments 1 and 2:

***Despite these local improvements, the overall impact of the updated land cover on domain-averaged SWE performance remains modest. This reflects the highly heterogeneous spatial influence of LC updates: while some vegetation transitions substantially reduce SWE errors, many regions either exhibit no LC change or experience transitions with limited influence on snow simulations. Consequently, domain-averaged statistics improve only slightly, despite clear local improvements in specific regions. These results also suggest that land cover representation is just one factor influencing the accuracy of SWE simulations. The remaining SWE errors are influenced by uncertainties in atmospheric forcing, snowfall, precipitation phase partitioning, snow physical parameterisations, vegetation structural characteristics and topographic effects. Therefore, further improvements in SWE simulations will likely require a combination of updated LC information and advances in atmospheric forcing, snow process representation and satellite data assimilation rather than relying on LC improvements alone.***

We also added this at the end of the Conclusion section:

***Although the updated land cover dataset improves SWE simulations slightly, remaining SWE errors suggest that land cover information alone is insufficient to resolve all model uncertainties. Further advances in atmospheric forcing, snow process representation and data assimilation will likely be required for future improvements.***

**No results on the performance of the model and the change of LC to simulate LST are discussed in the abstract. This should be added.**

**RESPONSE 1.3:**

We thank the reviewer for this suggestion. We agree that the abstract should also summarize the impact of the updated land cover on LST simulations. We have therefore added a sentence describing that the updated land cover has only a limited effect on daytime and nighttime LST performance, while reducing systematic SWE errors.

We added this to the abstract:

***The updated land cover has a very limited impact on the model's performance for land surface temperature, indicating that the impact of LC updates is more noticeable for SWE than for LST.***

**In section 2.1, I think it would be important to state here that the LST corresponds to the radiative surface temperature. It is stated later, but I think it could be introduced earlier**

**RESPONSE 1.4:**

We thank the reviewer for this suggestion. We agree that this information should be introduced earlier in the manuscript. We have therefore added a short description in Section 2.1 explaining that the simulated LST corresponds to the effective radiative surface temperature computed by the ISBA radiative transfer scheme. We moved the following paragraph from Section 3 to Section 2.1:

***In the ISBA model, the simulated LST represents the radiative surface temperature of the grid cell. This variable is not simply the average of the temperatures of the soil, vegetation and snow. Instead, it is computed within the ISBA radiative transfer scheme as an effective surface temperature, integrating the contributions of all surface components (i.e. soil, the vegetation canopy, and snow) according to their fractional coverage and radiative properties.***

**The authors now specify that a linear interpolation is used to upscale values to the 0.25o grid. I am not sure that a linear interpolation is the most appropriate interpolation. Can this be discussed?**

**RESPONSE 1.5:**

We thank the reviewer for this comment. Linear interpolation was selected because it preserves the large-scale spatial gradients while avoiding the introduction of artificial discontinuities when aggregating the higher-resolution satellite products to the 0.25° ISBA grid. Since our objective is the comparison at the model grid resolution rather than the preservation of sub-grid variability, linear interpolation represents an appropriate compromise. We have clarified this point in the manuscript in Section 2.2:

***Linear interpolation was selected because the aim of this study is to compare the ISBA model at its native resolution, rather than to preserve sub-grid variability.***

**I would suggest moving section 2.4 before the current section 2.3 ("ERA5 SWE and LST data")**

**RESPONSE 1.6:**
We thank the reviewer for this suggestion. We agree that presenting the observational datasets before the ERA5 reference improves the logical flow of the manuscript. We have therefore reordered Sections 2.3 and 2.4 accordingly.

**I think a deeper discussion or introduction on the differences between the LC maps is needed. Does the new LC correspond to changes of vegetation type over time or is it more accurate than the previous one due to improved algorithms/satellite data?**

**RESPONSE 1.7:**

We thank the reviewer for this important comment. We agree that the nature of the differences between the two land cover datasets should be clarified before presenting the comparisons. We have therefore expanded Section 2.5 to explain that ECOCLIMAP-SG is not an updated map reflecting vegetation changes over time. Rather, it combines more recent satellite observations (ESA CCI LC and Copernicus vegetation products), higher spatial resolution and improved classification methodologies. Consequently, the differences between Old LC and New LC reflect both genuine land-cover changes that occurred between the early 2000s and 2010, and improvements in the representation of vegetation resulting from newer datasets and algorithms. We propose to add this in Section 2.5, before the description of Figure 1:

***Consequently, the differences between the two datasets reflect actual land-cover changes that occurred between the early 2000s and 2010, as well as improvements resulting from more recent satellite observations and a higher spatial resolution. Therefore, the New LC dataset should not be interpreted as solely depicting land-cover change, but rather as providing a more realistic representation of land-cover conditions during the study period.***

**Figure 1c, could it be specified in this caption that the grid cells where ESA CCI SWE are available correspond to the blue ones?**

**RESPONSE 1.8:**

We thank the reviewer. The figure caption has been revised to explicitly state that the blue grid cells correspond to locations where valid ESA CCI SWE observations are available.

**L. 185-188: These sentences stating different periods are confusing. Why are simulations repeated four times over the 1981-1989 and 2010-2019 periods when the study period is 2010-2022?**

**RESPONSE 1.9:**

We thank the reviewer. We agree that the original wording could be confusing. The description of the spin-up procedure has been revised to clarify that the repeated simulations apply only to the spin-up period and not to the evaluation period (2010–2022). We changed L. 187 to:

***Specifically, a 200-year generic spin-up was followed by two multidecadal spin-ups for Old LC and New LC, as described by Liu et al. (2025). The multidecadal spin-ups consisted of the same 10-year period repeated at least four times to ensure equilibrium by 31 December 2009, prior to the start of the evaluation period (2010–2022).***

**As before, I think the quality of the graphs should be improved. I still find them a bit blurry in the PDF. Maybe increasing the dpi to 300 could help. In addition, the texts (labels, titles), could be bigger, as well as the ticks (degree values). In figures showing the time series (e.g. Fig. 3, Fig.8), the reference data ESA CCI could be put in black to highlight that it's the reference data**

**RESPONSE 1.10:**

We thank the reviewer for these suggestions. All figures have been regenerated at 300 dpi with larger font sizes for labels, titles and tick marks. In addition, the reference ESA CCI time series has been plotted in black in Figures 3 and 8 to facilitate comparison with the model simulations.

**Figure 7: consider doing a 2x2 plot**

RESPONSE 1.11:

Done.

**L.324-325: "Over 90 % of the domain has RMSD values below 40 mm, indicating generally low errors in snow estimation." I think the threshold of 40 mm is quite high as the snowpack is generally under 100 mm (Figure S7). Could the normalized RMSD be looked at?**

RESPONSE 1.12:

We thank the reviewer for this valuable suggestion. We agree that expressing the errors only in absolute units (mm) can be misleading because the climatological SWE varies substantially across Europe. Following the reviewer's recommendation, we computed the normalized RMSD (nRMSD), defined as the RMSD normalized by the mean ESA CCI SWE at each grid cell.
We have therefore replaced panel (c) of Figure 7, which previously showed the ubRMSD CDF, with the CDF of the normalized RMSD, providing a more meaningful assessment of the relative model error. In addition, the previous ubRMSD CDF has been moved to the Supplementary Material, where it is presented together with the CDF of the normalized ubRMSD (new Figure S7).
The new analysis confirms the conclusions obtained from the absolute metrics. ERA5 still exhibits the lowest relative errors over most of the domain, while the New LC simulation shows a systematic reduction in normalized RMSD compared with the Old LC, demonstrating that the improvement associated with the updated land cover remains evident even after accounting for the local climatological SWE magnitude. The corresponding discussion in Section 4.2 has been revised accordingly.
We replaced this paragraph, at the end of Section 4.2, by:

***Over 90% of the domain has RMSD values below 40 mm. As climatological SWE varies substantially across Europe, nRMSD (Fig. 7c) provides a more meaningful measure of relative error. This metric corroborates the findings derived from absolute RMSD, indicating that ERA5 exhibits the lowest relative errors across the majority of the domain. Conversely, the New LC simulation consistently shifts the nRMSD distribution towards lower values compared to Old LC, suggesting a notable reduction in relative SWE errors following land cover updates. While the bias spread of New LC is not reduced compared to Old LC, instances of systematic SWE overestimation are significantly reduced. Figure 7 also shows a slight improvement in the correlation CDF. The corresponding CDFs of ubRMSD and normalised ubRMSD are provided in the supplementary material (Fig. S7).***

**L. 351: 'DBF', 'CNF': consider writing them fully**

RESPONSE 1.13:

We thank the reviewer. The abbreviations 'DBF' and 'CNF' are now written out in full. This is also the case in the caption of Table 2.

**Figure 9: it is not clear to me if these maps are only for the winter seasons or over the whole years? I am wondering as this figure is used to relate the LST errors with the modelled SWE bias.**

**RESPONSE 1.14:**

We thank the reviewer. Figure 9 shows the mean daytime and night-time LST for the whole 2010-2022 study period, rather than just the winter season. The caption of Fig. 9 was updated accordingly.

## REVIEWER 2

**The authors have provided thorough and thoughtful responses. I appreciate the additional analysis around land cover changes (Section 5.1). The manuscript is clean, well-prepared and well-written. I have a few very minor notes where the old text doesn't fully reflect the updated figures and/or findings. Otherwise, the manuscript is suitable for publication.**

We thank the reviewer for the positive assessment of the revised manuscript and for the constructive minor comments. We have revised the abstract, the interpretation of the SWE performance results, Figure 8, and the conclusion to better reflect the updated figures and the new analysis of land-cover and sub-grid vegetation-fraction changes.

**Suggest adding a phrase touching on the impact of change in sub-pixel vegetation fraction changes also being important in SWE biases. See also comment L439.**

**RESPONSE 2.1:**

We thank the reviewer for this suggestion. We agree that the abstract should better reflect the new analysis presented in Section 5.1. We have therefore added a sentence indicating that changes in both dominant land-cover classes and sub-grid vegetation fractions contribute to changes in SWE bias. We added in the Abstract:

***Changes in SWE are also found in areas where the dominant LC remains unchanged. This indicates that modifications in sub-grid vegetation fractions can affect snow-vegetation interactions and SWE bias.***

**L275-277: Check the consistency of the revised sentence. If the New LC does not change the correlation how can it be concluded that this indicates a 'marginal improvement in temporal agreement'?**

**RESPONSE 2.2:**

We thank the reviewer for pointing out this inconsistency. We agree that if the Pearson correlation coefficient remains unchanged, the text should not refer to an improvement in temporal agreement. We have corrected the sentence in L.275 accordingly:

***Using the New LC does not change the mean Pearson correlation coefficient between the modelled and observed SWE (R = 0.84). This indicates that there is essentially no change in the temporal agreement with ESA CCI SWE.***

**L277-278: Maybe marginally reduced rather than notably? Reduction is on the order of ~2mm and in a relative sense, RMSE and uRMSD only reduced by ~4-5%. Larger relative MB reduction (~23%) due to smaller initial magnitude.**

**RESPONSE 2.3:**

We thank the reviewer for this helpful comment. We agree that the term "notably reduced" overstates the magnitude of the improvement in RMSD and ubRMSD, particularly after correcting the ESA CCI filtering procedure. We have therefore revised the wording to better reflect the quantitative results. The revised text now emphasizes that the updated land cover leads to modest

reductions in RMSD, ubRMSD and mean bias, while the temporal correlation with ESA CCI SWE remains essentially unchanged.
The revised paragraph, just before Table 1, now reads:

***Using the New LC does not change the mean Pearson correlation coefficient between the modelled and observed SWE (R = 0.84). This indicates that there is essentially no change in the temporal agreement with ESA CCI SWE. On the other hand, RMSD, ubRMSD, nRMSD, and MB are slightly reduced in the New LC simulations: from 34.7 mm to 32.9 mm, 34.1 mm to 32.6 mm, 62 % to 59 %, and 6.4 mm to 4.9 mm, respectively. This suggests that the more detailed and updated LC dataset more accurately represents snow accumulation processes in ISBA, while only achieving modest improvements in domain-averaged SWE performance. This is primarily due to reductions in systematic errors, rather than changes in temporal agreement with ESA CCI SWE.***

**L327-328: Text related to Figure 7. I found it hard to identify the reduced instances of systematic overestimation from the updated Figure 7d. Improvement in correlation isn't very clear in the updated Fig 7a.**

**RESPONSE 2.4:**

We thank the reviewer for this observation. We agree that the updated Figure 7 does not clearly show a visible improvement in the correlation CDF or a marked reduction in systematic overestimation. We have therefore softened this interpretation and now describe Figure 7 as showing broadly similar correlation distributions for Old LC and New LC, with only modest changes in the error distributions. We also updated Fig. 7 by adding the nRMSD CDF as it presents more differences between New LC and Old LC.
We reformulated the whole interpretation of Figure 7 as:

***ERA5 shows the highest R values and the lowest RMSD and ubRMSD values overall. Over half of the points have correlation values above 0.8, indicating that all three models (ERA5, Old LC and New LC) accurately depict the seasonal changes in SWE. However, the correlation distributions of Old LC and New LC are very similar, which is consistent with the domain-averaged statistics shown in Table 1. Over 90% of the domain has RMSD values below 40 mm. As climatological SWE varies substantially across Europe, nRMSD (Fig. 7c) provides a more meaningful measure of relative error. This metric corroborates the findings derived from absolute RMSD, indicating that ERA5 exhibits the lowest relative errors across the majority of the domain. Conversely, the New LC simulation consistently shifts the nRMSD distribution towards lower values compared to Old LC, suggesting a notable reduction in relative SWE errors following land cover updates. Although the bias spread of the New LC is not reduced compared to the old LC, instances of systematic SWE overestimation have been significantly reduced. Therefore, the main difference between the two ISBA simulations is not an obvious improvement in temporal correlation, but rather a slight reduction in RMSD, ubRMSD and mean bias, alongside a more notable reduction in nRMSD. The corresponding CDFs of ubRMSD and normalised ubRMSD are provided in the supplementary material (Fig. S7).***

**Figure 8: I find it hard to distinguish between the red and orange/yellow colours. Maybe try a darker red or a brighter or more distinct colour for Old LC. Check colour-blind friendly pallets.**

**RESPONSE 2.5:**

We thank the reviewer for this helpful suggestion. We agree that the original colour scheme made it difficult to distinguish between the Old LC and New LC simulations. This is not only because the colours were too similar, but also because the differences between the two simulations are generally very small. As a result, the two curves largely overlap throughout the time series, and one curve partially obscures the other over many periods. To improve readability, we have revised the colour scheme in Figure 8 by displaying the ESA CCI reference in black and using a darker red colour for Old LC and green for LST CCI product, providing a clearer distinction between the three datasets while maintaining a colour-blind-friendly palette. In addition, we introduced different marker symbols for each dataset, allowing the three time series to be more easily distinguished even where the curves overlap.

**L375-377: Near-duplicate sentences. Maybe select the second one?**

**RESPONSE 2.6:**

We thank the reviewer for noticing this duplication. We have removed the first duplicated sentence and retained the clearer second formulation.

***Overall, these results suggest that SWE does not systematically respond to all changes in LC. Instead, its response depends on the type of transition and the extent of the associated structural changes to the vegetation.***

**L439: Maybe add something that clearly states it is changes in both the LC and the sub-pixel vegetation? See Abstract comment**

**RESPONSE 2.7:**

We thank the reviewer for this suggestion. We have clarified in the conclusion that the SWE response is influenced not only by changes in dominant LC classes, but also by changes in sub-pixel vegetation fractions.

***Additionally, changes in snow water equivalent (SWE) are not confined to grid cells where the dominant land cover changes. They can also result from modifications in sub-pixel vegetation fractions, which affect canopy density, vegetation composition, and snow–canopy interactions.***