REVIEWER 1

COMMENT 1.1. In this manuscript, the authors are presenting different simulations of snow water equivalent (SWE) across Europe with the ISBA model using two different land cover databases. They compared their simulations against ERA5 snow analyses as well as against the ESA CCI SWE databases. They showed that their simulation with the newest land cover database provided on average better SWE estimates over the full domain, but still some discrepancies with ERA5 or the SWE database could be observed. I found the paper well written and clear. I think this manuscript is fit for publication after a few improvements. Additional analysis is needed to examine how land cover changes between the two databases affect SWE estimates, which may explain why performance has declined in regions like Northern Europe.

RESPONSE 1.1:

Thank you for your constructive comments. We have attempted to address each one of them below. In particular, we propose adding extra material to make the effect of land cover on SWE simulations clearer.

COMMENT 1.2. Major comment: I think additional analysis should be done to explain the degradation in SWE in norther Europe (Figs. 4, 5, and L. 243-246). I think the degradation should be put in perspective with the new land cover in this area from temperate grasslands to bare soil with no vegetation (Fig. 1). Could this degradation be caused by a poor representation of certain processes in this new land classification?

RESPONSE 1.2:

We thank the reviewer for this insightful comment and for suggesting a potential physical explanation for the degradation observed in Northern Europe. We would like to clarify that the SWE score degradation observed in Northern Europe is mainly concentrated in Sweden and Finland, whereas the transition from temperate grasslands to bare soil, as mentioned by the reviewer, occurs in mountainous grid cells that are not included in the SWE evaluation. All results (see Figures 5 and 6 in the manuscript) are computed using the ESA CCI SWE data mask, which excludes regions without reliable CCI SWE values. Consequently, areas where such transitions may occur are not included in the evaluation. We reviewed the procedure for filtering ESA CCI SWE data and found an inconsistency when comparing modelled and observed SWE values for Old LC and New LC (see Fig. R1.1c). This inconsistency is clearly visible in Fig. R1.1c. Figure R1.2 shows the scatterplots again, but with the revised procedure. In the initial procedure, the statistics were not always calculated over the same set of grid cells and dates with valid ESA CCI SWE data. This introduced artefacts in regions with sparse or intermittent snow data. In the revised procedure, all statistics are computed only where ESA CCI SWE is available, after the datasets have been temporally aligned and restricted to snow-relevant grid cells (i.e. grid cells with valid snow data). Using the revised procedure for filtering SWE data ensures the use of valid SWE data and strict temporal alignment between modelled and observed values. This improves the agreement between modelled SWE and CCI SWE, increasing the R value from around 0.7 to 0.84, as can be seen in Table R1.1, which is a revised version of Table 1 in the original manuscript. Table R1.1 also shows that the mean bias has been significantly reduced. Figure 3 in the current manuscript will be updated accordingly.

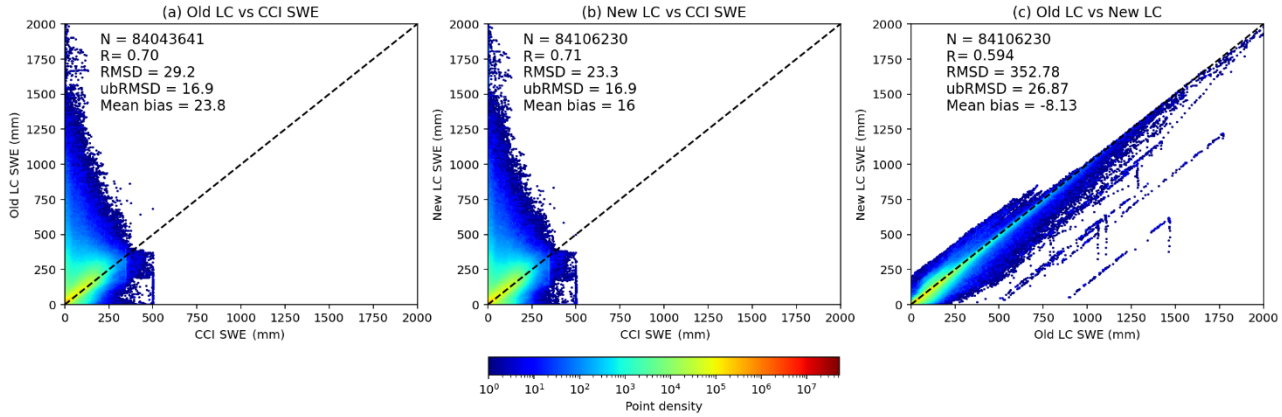

Figure R1.1*. SWE scatterplots: (a) Old LC vs. CCI SWE, (b) New LC vs. CCI SWE, (c) Old LC vs. New LC.*

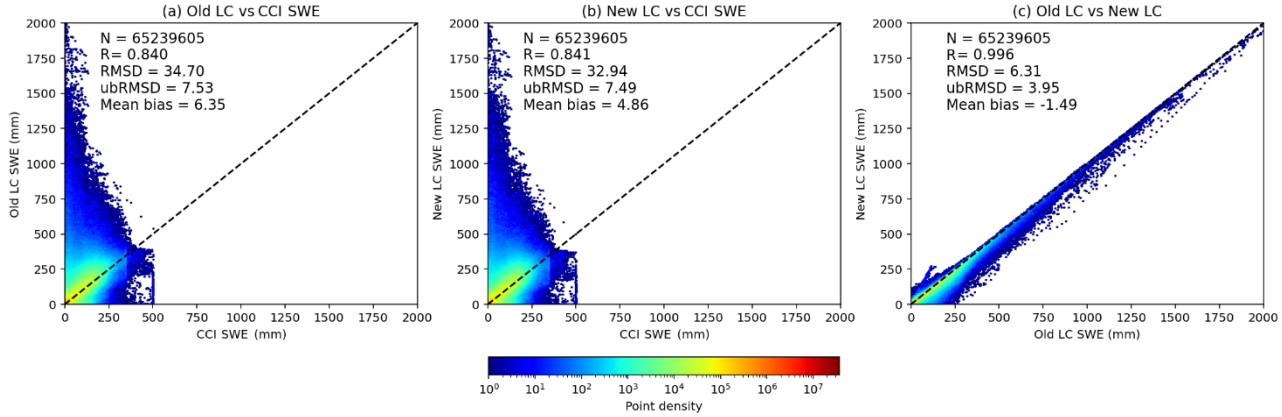

Figure R1.2*. Same as Fig. R1.1, except for the revised procedure for filtering ESA CCI SWE data.*

Table R1.1. Summary Mean grid-cell level score values of the New LC and Old LC simulations for SWE and LST (both daytime and nighttime) over the 2010–2022 period. SWE score values are over the whole European domain. LST score values are for the westernmost part of the domain (10°W-30°E, 28.125°N-71.875°N). The number of data and score values for ERA5 are also shown**.**

Model	**vs. CCI variable**	***R***	**RMSD**	**ubRMSD**	**MB**	**Number**
ERA5	SWE (mm)	0.90	23.4	23.4	0.9	65,199,450
Old LC		0.84	34.7	34.1	6.4	65,239,605
New LC		0.84	32.9	32.6	4.9	65,239,605
ERA5	Daytime LST (K)	0.96	4.8	4.4	- 1.8	18,056,129
Old LC		0.96	6.0	4.8	-3.6	19,311,176
New LC		0.96	6.1	4.9	- 3.7	19,311,164
ERA5	Nighttime LST (K)	0.97	3.0	2.5	1.6	23,109,907
Old LC		0.95	3.2	3.1	0.8	24,259,011
New LC		0.95	3.1	3.0	0.8	24,274,817

We revised Fig. 5 in light of the change in filtering strategy. As can be seen in Fig. R1.3 (a revised version of Fig. 5), the impact of this correction is significant. Although the previously reported degradation in Northern Europe has been reduced, it is still present. To investigate the slight degradation of *R* and RMSD observed in Sweden and Finland, we plotted the grid cells in which the dominant PFT changed in Fig. R1.4. It appears that the dominant PFT remained largely unchanged in these regions. Conversely, significant changes occurred in the other PFTs, as illustrated in Fig. R1.5. When analysed together, Figures R1.3, R1.4 and R1.5 reveal a consistent spatial pattern. Regions where the dominant PFT changes (Fig. R1.4) correspond to areas with the greatest impact on SWE (Fig. R1.3). Meanwhile, regions with non-dominant PFT changes (Fig. R1.5) exhibit much weaker SWE differences. This behaviour is particularly evident in Sweden and Finland, where the slight score degradation observed in Fig. R1.3 occurs in areas with no dominant PFT transition but only small sub-grid vegetation changes. The small magnitude of the sum of absolute differences in PFT (Σ|ΔPATCH|) in these regions is consistent with the weak amplitude of the SWE degradation. Focusing on grid cells in Fig. R1.3 in Sweden and Finland that exhibit a RMSD degradation when New LC is used, we find that only a limited proportion of the snow-relevant grid cells (23.4%) exhibit a change in dominant PFT, mostly located in northern Finland, while the majority (76.6%) remain unchanged. However, significant sub-grid vegetation redistribution occurs within these 'unchanged' grid cells, with a median value of Σ|ΔPATCH| of ~0.36 and over than 21% of pixels exceeding 0.5. In terms of the SWE response, grid cells with a change in the dominant PFT show a stronger signal (mean ΔSWE ~-4.3 mm), whereas pixels without a dominant change exhibit a weaker decrease (mean ΔSWE ~-1.7 mm). This suggests that the slight degradation is primarily linked to moderate sub-grid vegetation adjustments rather than widespread changes in dominant land cover. Overall, this analysis suggests that the initial reported degradation was largely caused by inconsistent filtering, and that the remaining differences are minor and mainly related to small sub-grid vegetation adjustments rather than significant changes in the dominant land cover.

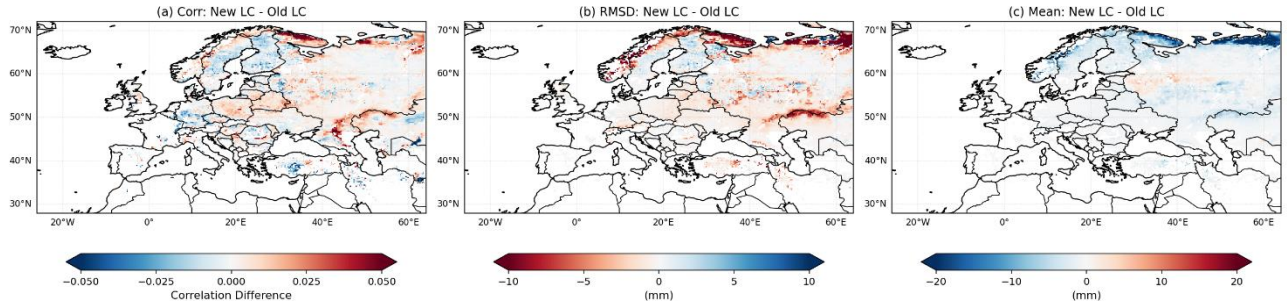

Figure R1.3*. (Revised Fig. 5) Spatial differences in the statistical performance metrics of New LC and Old LC SWE simulations with respect to the ESA CCI SWE product over the whole European domain, over the 2010–2022 period: difference in (a) R, (b) RMSD and (c) mean difference of SWE between the two simulations. For R and RMSD, red zones indicate improvement (e.g. higher correlations or 240 lower errors) when using New LC, while blue zones indicate degradation.*

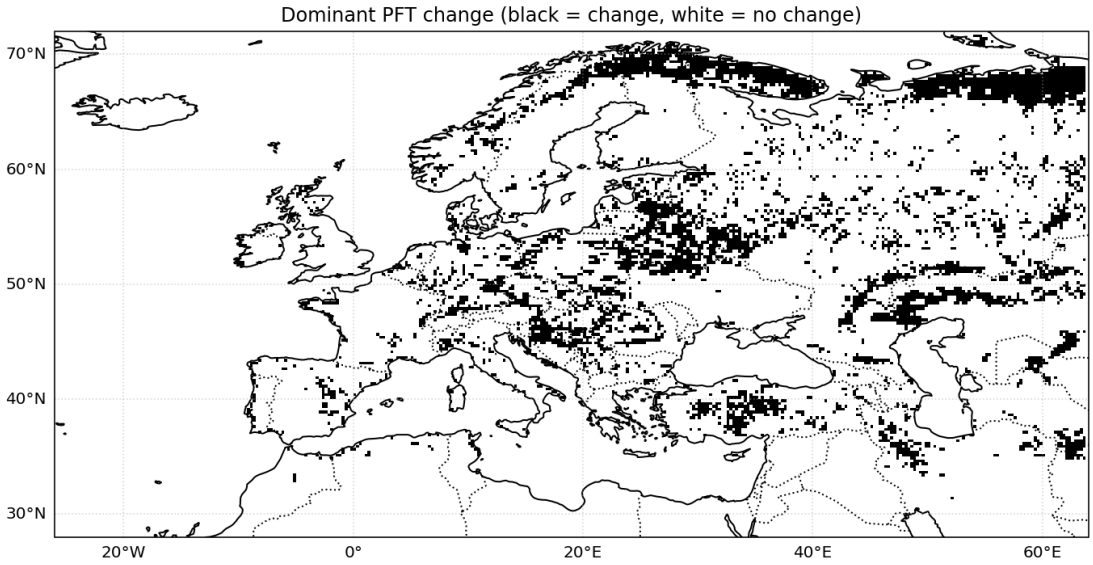

***Figure R1.4**. Spatial distribution of dominant PFT changes between New and Old LC (ECOCLIMAP-SG and ECOCLIMAP-II, respectively) over snow-relevant grid cells. Black grid cells indicate a change in dominant PFT, while white ones indicate no change.*

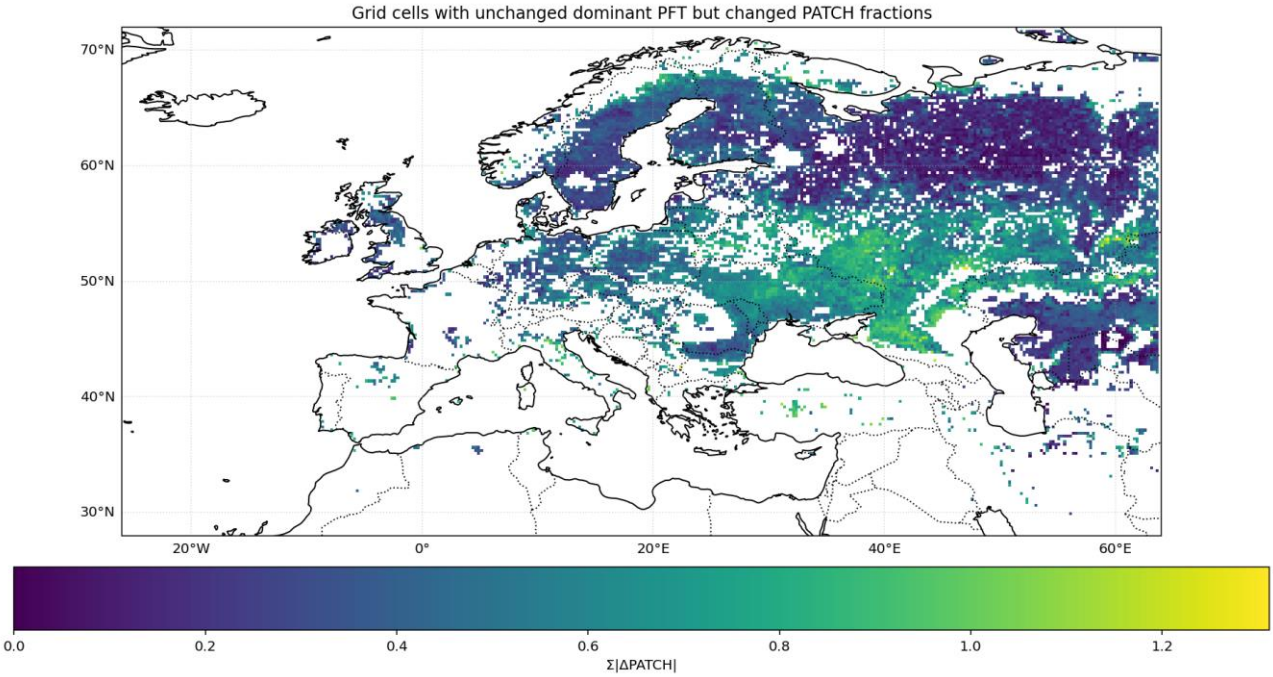

***Figure R1.5**. Spatial distribution of grid cells with unchanged dominant PFT but modified sub-grid vegetation fractions (PATCH). The colour scale represents the sum of absolute changes in PFT fractions ($\Sigma|\Delta PATCH|$). Values may exceed 1 due to the use of absolute differences across multiple PFT fractions.*

COMMENT 1.3. Major comment: Section 2.1 - I think it would be interesting to have a more complete description of the ISBA model, including the interaction of the snow with the atmosphere, the forest, and the ground. I don't find the description of ISBA-A-gs pertinent for this study.

RESPONSE 1.3:

We thank the reviewer for this insightful comment. We agree that the description of the ISBA model in Section 2.1 was insufficiently detailed, especially with regard to the interactions between snow, vegetation, soil and the atmosphere. We propose revising Section 2.1 to provide a more complete and physically consistent description of the ISBA model, particularly with regard to the representation of snow processes. We also intend to enhance the description of vegetation processes in the revised manuscript, including the ISBA-A-gs configuration, to illustrate more clearly how vegetation dynamics influence snow accumulation, interception and melt processes. These additions will justify the relevance of the model configuration for this study, which focuses on snow simulations and their sensitivity to land cover, more effectively. In this study, the representation of leaf area index (LAI) is particularly important, as it directly affects the surface energy balance and snow processes. Vegetation influences snow accumulation, sublimation and melt dynamics by controlling canopy radiative transfer, turbulent heat fluxes and interception capacity. Dense canopies can, for instance, intercept snowfall, thereby reducing the amount that reaches the ground. Therefore, interactive simulation of LAI is essential to capture the impact of changes in land cover on SWE and land surface temperature (LST), even in the absence of direct observational constraints. The ISBA-ES (Explicit Snow) scheme (Boone and Etchevers, 2001; Decharme et al., 2016) is employed for the representation of snow processes. It resolves the vertical evolution of the snowpack through multiple layers. It simulates processes such as snow accumulation, compaction, metamorphism, melting, refreezing and sublimation, as well as heat exchanges between the snowpack, soil, vegetation canopy and atmosphere.

COMMENT 1.4. Major comment: L. 96, 113, 125. It is mentioned that the product is regridded/interpolated/resampled to a coarse resolution. What kind of regridded/interpolation method is used?

RESPONSE 1.4:

The ERA5 skin temperature data was originally provided on a regular 0.25° × 0.25° latitude–longitude grid. Therefore, no change in spatial resolution was applied. The fields were subset spatially to the study domain and only interpolation was used to ensure exact alignment with the target 0.25° analysis grid. The ESA CCI SWE and LST products were regridded to the target 0.25° regular latitude–longitude grid using linear interpolation (SciPy griddata implementation). For LST, this involved changing the resolution from 0.05° to 0.25°, whereas for SWE, the resolution was adjusted from 0.1° to 0.25°. Regridding was applied independently at each time step to ensure spatial consistency between the datasets.

COMMENT 1.5. Major comment: The modelled surface temperature is compared to observations (CCI-LST). The authors need to explain how they calculated grid-averaged surface temperatures (between soil, vegetation, snow, …) that they compared to CCI-LST. Also, this evaluation of LST should maybe be mentioned in the title of the manuscript and in the abstract. As it is right now, this evaluation feels out of place in the manuscript.

RESPONSE 1.5:

We thank the reviewer for this important remark. In the ISBA model, the surface temperature used for comparison with the LST derived from satellites corresponds to the diagnostic variable TSRAD_ISBA, representing the radiative surface temperature of the grid cell. This variable is not simply the arithmetic average of soil, vegetation or snow temperatures. 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, vegetation canopy and snow) according to their fractional cover and radiative properties. More specifically, TSRAD_ISBA is derived from the surface energy balance and corresponds to the temperature that would reproduce the longwave radiation emitted by the heterogeneous surface. As such, it can be directly compared with satellite LST products, which also represent a radiometric surface temperature. In our analysis, we extract TSRAD_ISBA at the model time steps that are closest to the times of the satellite overpasses (12:00 LT for daytime and 00:00 LT for night-time), thereby ensuring temporal consistency with the ESA CCI LST data. This will be clarified in the revised manuscript.

COMMENT 1.6. Major comment: L. 171-172 - why was the comparison of the surface temperature done in a sub-domain? Can this subdomain be shown on Fig. 1?

RESPONSE 1.6:

Thank you for your comment. The LST comparison was performed over a restricted subdomain to minimise the difference in time between the satellite data and the model output. AQUA MODIS LST corresponds to approximately 12:00 and 00:00 local solar time, whereas ISBA outputs are available every three hours in UTC. Over a large longitudinal extent, the local time corresponding to a given UTC output can vary significantly, which increases representativeness errors and introduces artificial differences that are unrelated to the physics of the model. Restricting the longitude range (10°W–30°E) reduces variability in local time across the domain, thereby improving the temporal consistency of the model–satellite comparison. The selected subdomain is indicated in Fig. 1 and illustrated in Fig. R1.6.

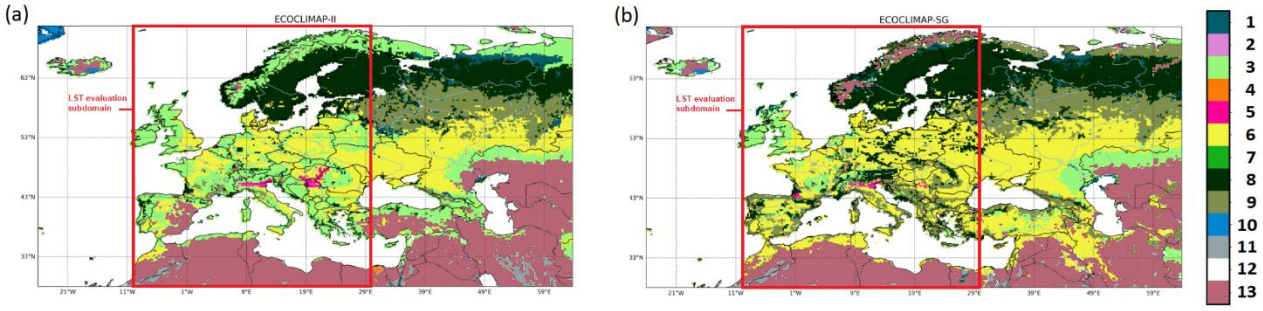

Figure R1.6*. Dominant land cover type over Europe at a spatial resolution of 0.25° x 0.25° as derived from (a) Old LC ECOCLIMAP-II (Faroux et al., 2013), (b) New LC ECOCLIMAP-SG (Calvet and Champeaux, 2020), with CCI LC v2.0.7 2010. The 12 dominants land cover types are indicated in the colour bar: 1 - flooded shrubs or grass, 2 - tropical grasslands, 3 - temperate grasslands, 4 - flooded trees, 5 - C4 crops (e.g. maize), 6 - C3 crops (e.g. wheat), 7 - broadleaf evergreen trees, 8 - coniferous trees, 9 - deciduous broadleaf trees, 10 - permanent snow and ice, 11 –rocks, urban, 12 – ocean and water bodies, 13 – bare soil with no 145 vegetation. The subdomain used for the LST evaluation is indicated by the red solid line.*

COMMENT 1.7. Minor comment: L. 160-161: it is not clear to me how the LAI is calculated, and it should be explained why it matters in this study.

RESPONSE 1.7:

We thank the reviewer for this comment. We agree that the original manuscript's description of LAI was not clear enough. To address this, we will revise Section 2.1 to explicitly explain how LAI is calculated in the ISBA-A-gs configuration. LAI is a prognostic variable derived from carbon fluxes that evolves dynamically in response to photosynthesis, leaf respiration and senescence processes. LAI is important in this study because it controls canopy radiative transfer, turbulent fluxes and snow interception, thereby directly influencing snow accumulation and melt processes, as well as land surface temperature. Therefore, an interactive representation of LAI is essential for assessing the impact of land cover changes on SWE and LST.

COMMENT 1.8. Minor comment: L. 187 - How are the anomalies calculated? The LAI anomalies also need to be further explained.

RESPONSE 1.8:

Thank you for this comment. Anomalies are computed as standardised anomalies (z-scores) with respect to a multi-year climatology. The z-score definition will be introduced explicitly in the 'Methods' section of the revised manuscript. For each grid point, the time series is grouped by calendar day (day of year), and the climatological mean and standard deviation are computed for each corresponding day of year across all available years. The scaled anomaly of a variable x at time t is then defined as follows:

$$x'(t) = \frac{x(t) - \mu_{DOY}(t)}{\sigma_{DOY}(t)}$$

where $x'(t)$ is the standardized anomaly at time t, and $\mu_{DOY}(t)$ and $\sigma_{DOY}(t)$ are the climatological mean and standard deviation of x for the corresponding day-of-year (DOY) at that grid point. This approach removes the seasonal cycle. The scaled anomalies are expressed in terms of standard deviations, allowing for consistent comparisons across regions and variables. For LAI and the other variables analysed in this study, the interpretation of the scaled anomalies presented in Figure 2 of the manuscript can be developed further to highlight the relationships and covariance between variables. The interpretation of the Hovmöller diagrams in Fig. R1.7 will be expanded upon in the revised manuscript. A notable feature is the winter 2020 period, when pronounced positive anomalies were observed simultaneously in LAI, skin temperature and SWE over high northern latitudes. This coherent signal highlights a large-scale warm anomaly affecting both vegetation activity and surface thermal conditions. Although SWE anomalies primarily occur in northern regions during winter, their co-occurrence with positive temperature anomalies during this event indicates strong coupling between vegetation dynamics and the surface energy balance in the presence of anomalously warm winter conditions.

COMMENT 1.9. Minor comment: L. 204 - ‘narrowing the gap with the observations”. I think another term than ‘gap’ should be used here. Bias? And I am not sure the CCI SWE product should be called ‘observations’?

RESPONSE 1.9:

We thank the reviewer for this comment. We agree that the term 'gap' was not precise enough, and that referring to the CCI SWE product as 'observations' could be misleading. To improve clarity and accuracy, we propose replacing “gap” with “difference” and “observations” with “ESA CCI SWE product” throughout the manuscript.

COMMENT 1.10. Minor comment: L. 205 - ‘discrepancies’. This should be explicitly explained in the manuscript rather than leaving it for the reader to interpret.

RESPONSE 1.10:

We thank the reviewer for this comment. We agree that the term 'discrepancies' was not sufficiently explicit. To improve clarity and avoid ambiguity for the reader, we propose replacing “discrepancies” with “differences” in the manuscript.

COMMENT 1.11. Minor comment: The quality of some figures (e.g. Fig. 2) should be improved.

RESPONSE 1.11:

We thank the reviewer for this comment. The quality of several figures, including Fig. 2, will be improved throughout the manuscript. In particular, we will enhance the resolution, readability and visual clarity of the figures. An updated version of Figure 2 is provided below (Figure R1.7) to illustrate this.

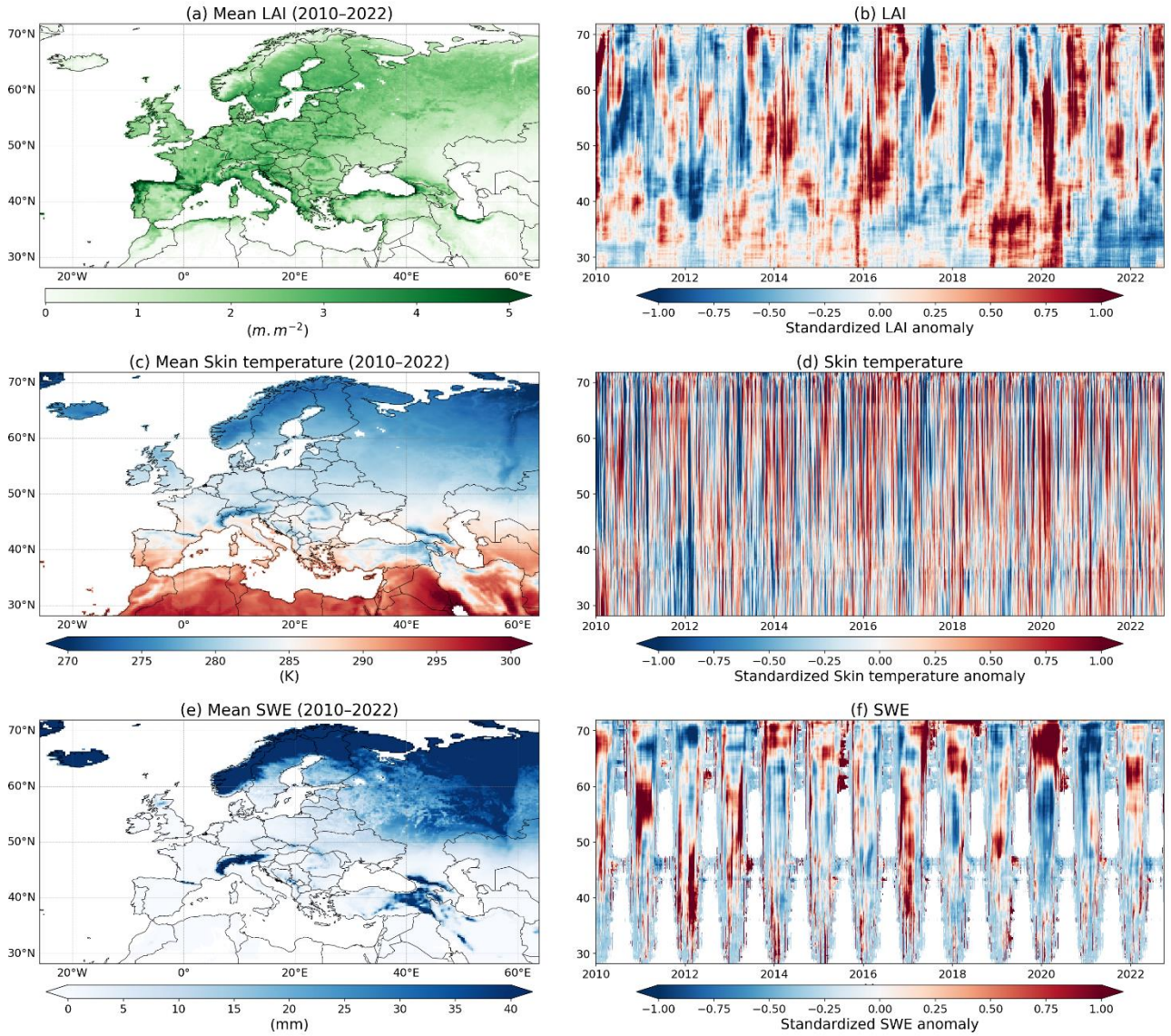

Figure R1.7*. New LC simulations over the whole European domain forced by ERA5 atmospheric variables from 2010 to 2022 at a 195 spatial resolution of 0.25 degree x 0.25 degree: Mean values (a, c, e) and Hovmöller plot (b, d, f) of scaled anomalies (z-score) of (a, b) LAI, (c, d) LST, and (e, f) SWE.*

COMMENT 1.12. Minor comment: L. 61 - I am surprised to see some big correlation differences in some areas that receive almost not snow, such as south-west of France, central Spain, and the UK. Maybe some filtering should be done?

RESPONSE 1.12:

We thank the reviewer for this pertinent comment. As discussed in response to point 1.2, we revisited the filtering applied when computing the statistical metrics and implemented a more consistent masking strategy based on observational availability, temporal alignment and snow-related conditions. Thanks to this improved filtering, the artefacts that were previously identified in regions with little or no snow cover (e.g. south-western France, central Spain and the UK) are no longer present. The revised correlation patterns are now more physically consistent, with significant differences mainly confined to snow-affected regions, particularly in northern and eastern Europe. The updated figure (Fig. R1.3) will replace Fig. 5 in the revised manuscript. Additionally, in the

ERA5 comparison (Fig. 6 in the manuscript), only the correlation panel was affected by this issue. This will be corrected and updated in the revised version and is presented below in Fig. R1.8.

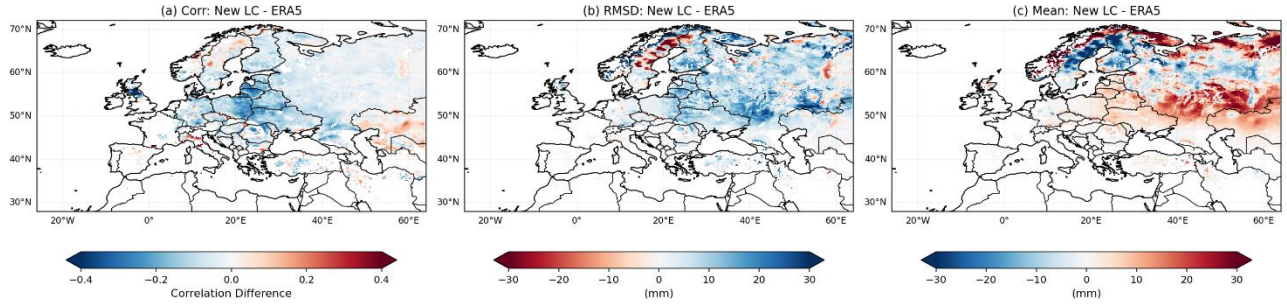

Figure R1.8. *Same as in Fig. R1.3, except for differences between New LC and ERA5.*