EDITOR – Dr. Nora Helbig

COMMENT. Thank you for your thorough author comments. The two referees raised several major issues that should be addressed in the revised manuscript. I read that you plan to address these accordingly, for example through an additional analysis examining changes in SWE with changes in LC, and an extended description of the ISBA model. Correcting newly found inconsistencies in the data filtering improved the modeled SWE degradation in Northern Europe. Given the major issues raised by the referees and the substantial revisions planned, including new figures and tables, I will send the revised manuscript out for further review.

RESPONSE:

Thank you for your feedback. We have redone the comparisons and prepared a revised version of the manuscript, which includes substantial changes to the results and discussion sections. We have also taken all of the suggestions made by the two reviewers into account in order to improve the quality of this work:

- Title was changed.
- The Abstract was updated.
- Section 2.1 was revised.
- Section 3 was extended in order to clarify the model evaluation protocol.
- All results were revised.
- A new subsection was added in Section 5, on the impact of LC changes on SWE.
- We added a paragraph to the Conclusion section.
- Figures 1, 2, 3, 5, 6, 7, and 8 were revised.
- New Figures were added to the Supplement: New Figs. S1, S2, S3, S4, S5, S7. Old Fig. S1 is now Fig. S6.
- Table 1 was revised. A new Table 2 was added.

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 addressed each one of them below. In particular, we added 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 ensured the use of valid SWE data and strict temporal alignment between modelled and observed values. This improved 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 was updated accordingly.

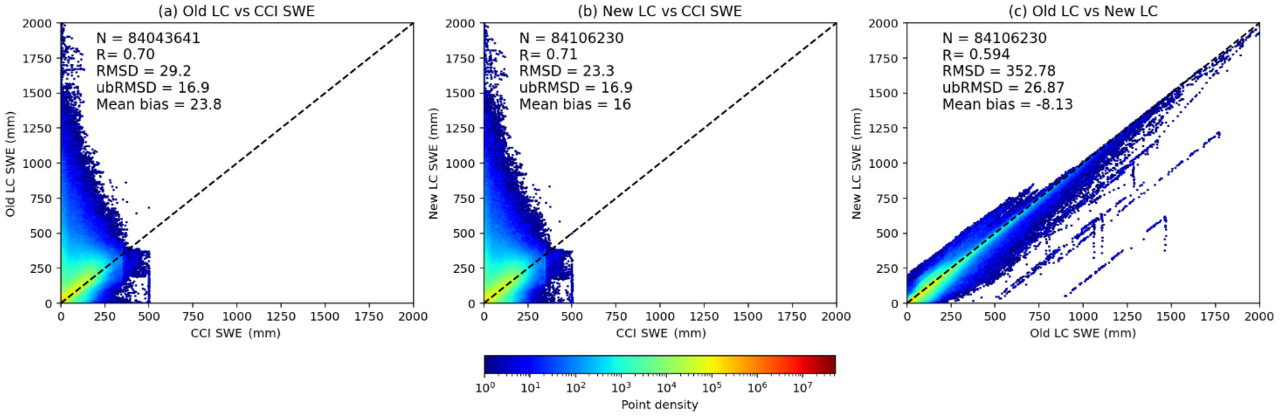

Figure R1.1. *SWE scatterplots: (a) Old LC vs. CCI SWE, (b) New LC vs. CCI SWE, (c) Old LC vs. New LC.*

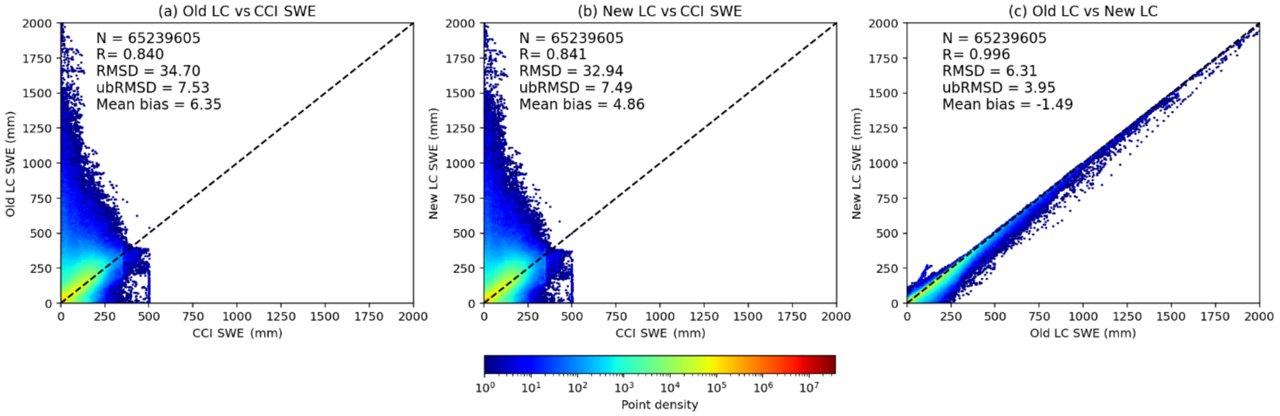

Figure R1.2 (New Figure S1). *Same as Fig. R1.1, except for the revised procedure for filtering ESA CCI SWE data.*

Table R1.1 (Revised Table 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	16,805,668
New LC		0.96	6.1	4.9	- 3.6	16,805,668
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	21,575,118
New LC		0.95	3.1	3.0	0.8	21,575,118

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 non-dominant PFT changes 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 non-dominant PFT changes 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.

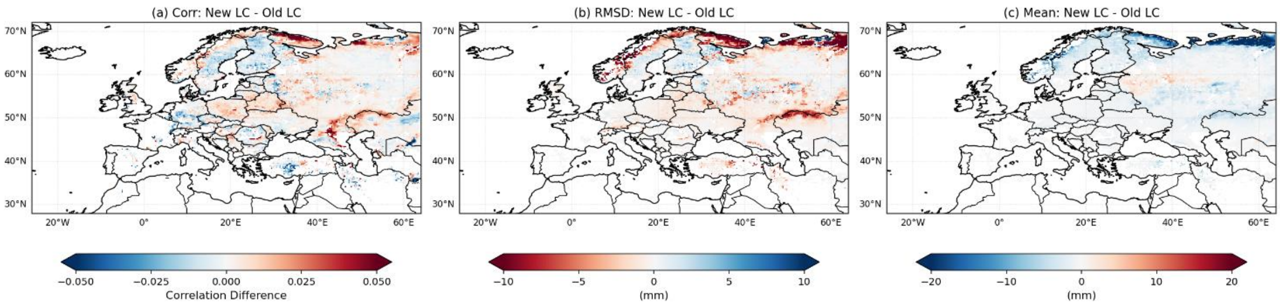

Figure R1.3. (***Revised Fig. 5***) *Spatial differences in the statistical performance metrics of New LC and Old LC SWE simulations with respect to the ESA CCI SWE product over the whole European domain, over the 2010–2022 period: difference in (a) R, (b) RMSD and (c) mean difference of SWE between 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.*

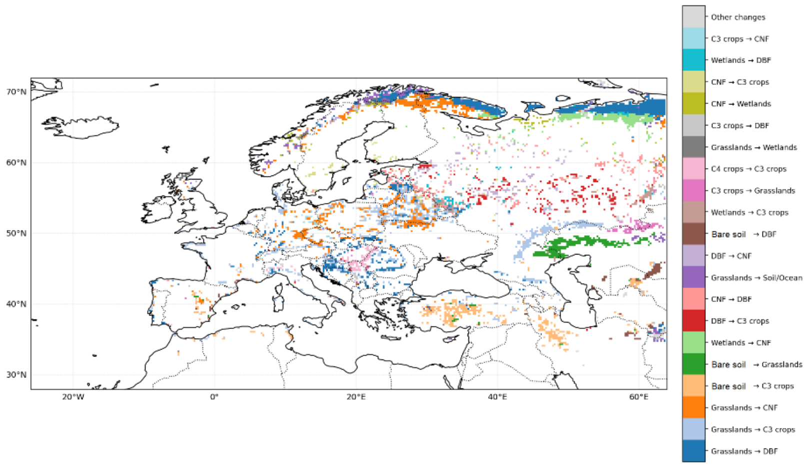

Figure R1.4 (New Fig. S4). *Spatial distribution of the top 20 most frequent dominant land cover transitions from Old LC to New LC over snow relevant grid cells where valid ESA CCI SWE data are available. All remaining transitions are grouped as "Other changes" Only transitions involving a change in dominant PFT are shown.*

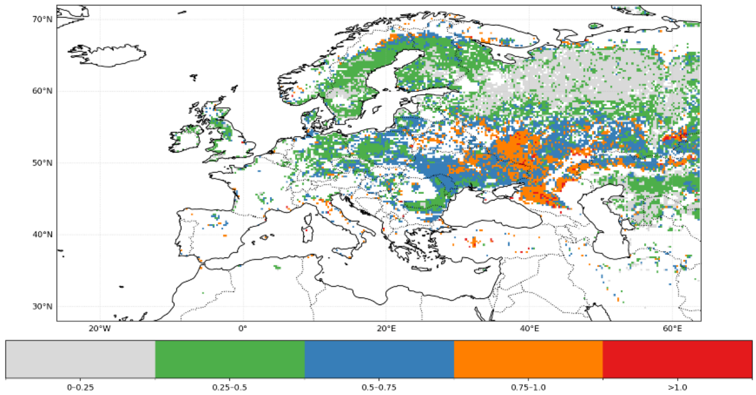

Figure R1.5 (New Fig. S5). *Spatial distribution of grid cells with unchanged dominant PFT but modified sub-grid vegetation fractions. The colour classes represent intervals of the sum of absolute changes in PFT fractions. Values may exceed 1 due to the use of absolute difference s across multiple PFT 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.
In the revised manuscript:

- The title was changed as: “Assessing the effect of land cover on ISBA snow water equivalent and land surface temperature simulations over Europe”
- In Section 2, it was mentioned that linear interpolation was used to adapt the CCI data to the model grid.
- Section 3 was enhanced.

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 13:30 and 01:30 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.

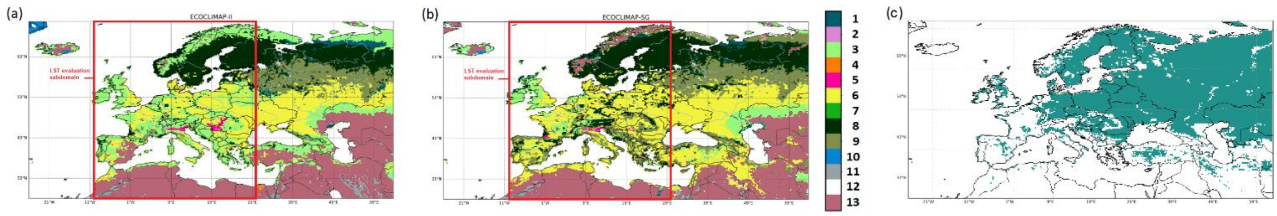

Figure R1.6 (Revised Fig. 1). *Dominant land cover type over Europe at a spatial resolution of 0.25° x 0.25° as derived from (a) Old LC ECOCLIMAP-II (Faroux et al., 2013), (b) New LC ECOCLIMAP-SG (Calvet and Champeaux, 2020), with CCI LC v2.0.7 2010, and (c) model grid cells where ESA CCI SWE is available. The subdomain used for the LST evaluation is indicated by the red solid line. 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 vegetation.*

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.

- Section 2.1 was revised accordingly.

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.
- Section 3 was revised accordingly.

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 replaced “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 replaced “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, was improved throughout the manuscript. In particular, we enhanced the resolution, readability and visual clarity of the figures. An updated version of Figure 2 is provided below (Figure R1.7) to illustrate this.

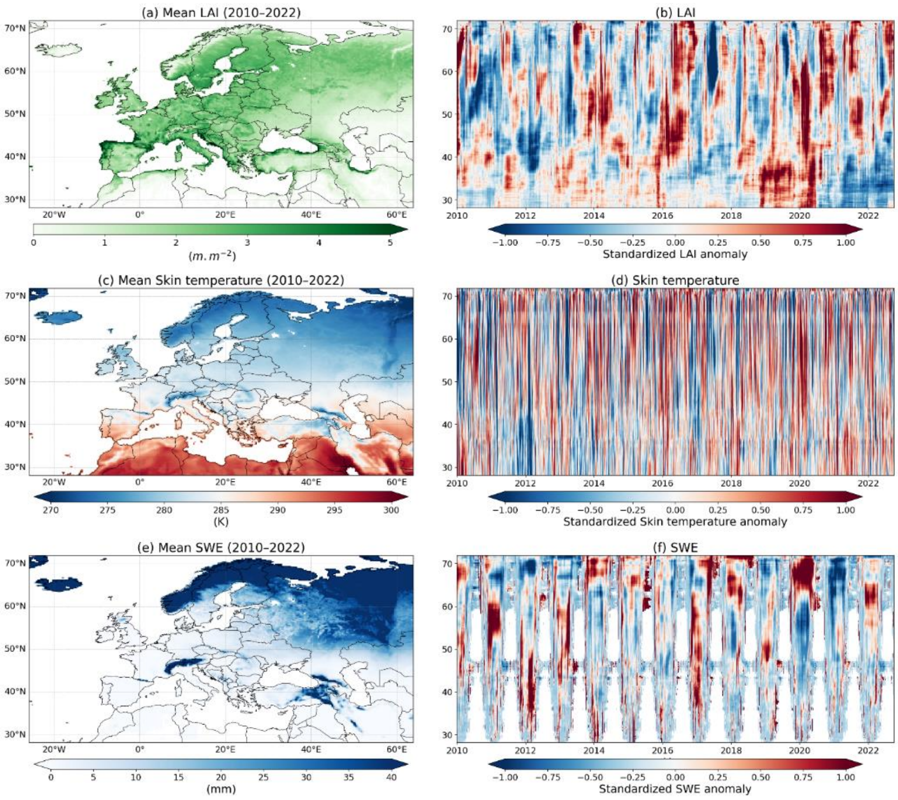

Figure R1.7 (Revised Fig. 2)*. New LC simulations over the whole European domain forced by ERA5 atmospheric variables from 2010 to 2022 at a 195 spatial resolution of 0.25 degree x 0.25 degree: Mean values (a, c, e) and Hovmöller plot (b, d, f) of scaled anomalies (z-score) of (a, b) LAI, (c, d) LST, 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) replaced 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 was corrected and updated in the revised version and is presented below in Fig. R1.8. Figure 7 was revised.

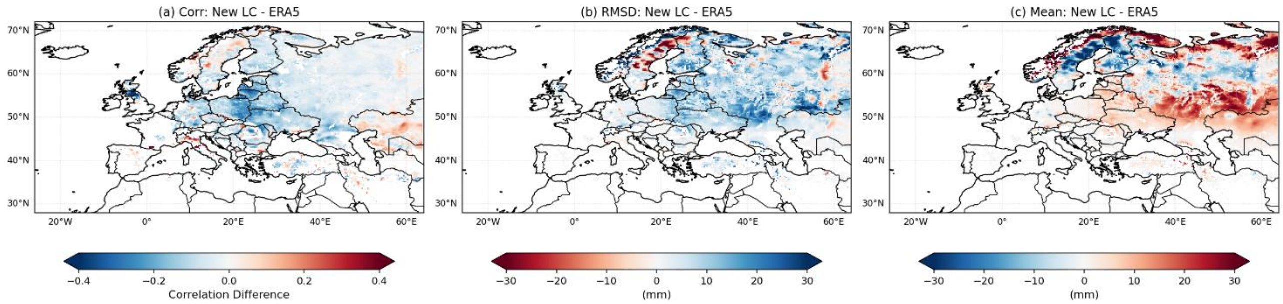

Figure R1.8 (Revised Fig; 6). Same as in Fig. R1.3, except for differences between New LC and ERA5.

REVIEWER 2

COMMENT 2.1. This investigates the impact of changing the land cover (LC) classification within ISBA (driven by ERA5 forcing) on the SWE estimates. ESA CCI SWE is used as the reference SWE dataset. Comparisons are also made with ERA5. Replacing the old LC information with one based on the ESA CCI LC v2.0.7 improved the SWE bias when compared to ESA CCI over the full temporal and spatial domain. Although New LC improved versus CCI, ERA5 often has better performance with CCI compared to the ISBA simulation. The authors also look at the impact of change in LC on LST but are unable to attribute LST differences to the changes in SWE. The manuscript is clear and well presented however I found the analysis and the discussion to be a bit limited. I would like to see additional analysis that looks at changes in SWE according to changes in LC classification. As it stands, the analysis shows that changing the LC information alters the SWE which is an expected result but it stops short of linking the SWE changes to specific changes in LC classification. Overall, while the manuscript is well-written and clear I found it lacking in analysis to support the main research objective.

RESPONSE 2.1:

We would like to thank the reviewer for their constructive and insightful comments. In response, we have conducted further analyses that explicitly link changes in snow water equivalent (SWE) to land cover (LC) transitions. These analyses include:
(1) the spatial identification of dominant land-cover transitions;
(2) the quantification of their frequency and impact; and
(3) the attribution of SWE and RMSD changes to specific transitions. The results are presented below in the form of new figures and a summarised transition table.

COMMENT 2.2. Did all LC changes lead to SWE changes? If not, which LC changes did not result in a SWE change (i.e. SWE agnostic to those different LCs). One suggestion is to quantify the number of pixels that experienced a reclassification and to further categorize those according to type of change (i.e. % of pixels that changed from 13 to 7) and finally to examine the change in SWE according to these LC changes.

RESPONSE 2.2:

We appreciate this relevant question from the reviewer. Following this recommendation, we conducted a thorough quantification of LC changes across the domain. Specifically, we calculated the number of grid cells affected by each transition and their relative spatial coverage (percentage of grid cells), as well as the associated changes in SWE and model performance metrics (bias, RMSD and ubRMSD). Our analysis shows that LC changes are highly heterogeneous in both space and impact. The most frequent transitions (e.g. grasslands to DBF and grasslands to CNF) affect approximately 4–5% of grid cells, whereas the majority of transitions affect less than 1% of the domain individually (see Table R2.1). While not all areas experiencing an LC change lead to a significant change in SWE, some of these transitions can produce noticeable local changes in SWE despite their limited spatial extent. This is illustrated by the scatterplot in Fig. R2.1, which compares SWE values simulated with Old LC (ECOCLIMAP-II) and New LC (ECOCLIMAP-SG). The very high correlation (R = 0.996) indicates that, at the domain scale, most grid cell values remain close to the 1:1 relationship. This means that a large proportion of LC changes do not result in significant differences in SWE. This behaviour is consistent with the spatial analysis. As Fig. R2.2 shows, only a subset of the snow-relevant domain exhibits a change in dominant plant functional type (PFT), while large areas remain unchanged. Furthermore, Fig. R2.2 and Table R2.1 show that although multiple LC transitions occur, their spatial extent and frequency vary considerably. The largest impacts, however, are associated with transitions from grasslands to forested classes (DBF and CNF), as these substantially modify vegetation structure and snow–vegetation interactions. These changes affect processes such as canopy interception, sublimation and radiative exchanges, resulting in more pronounced differences in SWE. Importantly, Fig. R2.3 shows that changes in SWE are not restricted to areas where the dominant PFT changes. Significant variations can also occur where the dominant PFT remains unchanged but sub-grid vegetation fractions are redistributed. The non-dominant PFT change fraction reflects the magnitude of these internal changes in vegetation composition. In such cases, modifications in canopy density or vegetation mixture can alter snow-related processes and induce SWE changes independently of dominant LC transitions. Overall, these results suggest that SWE does not systematically respond to all LC changes. Rather, its response depends on the type of transition and the magnitude of the associated structural changes in vegetation.

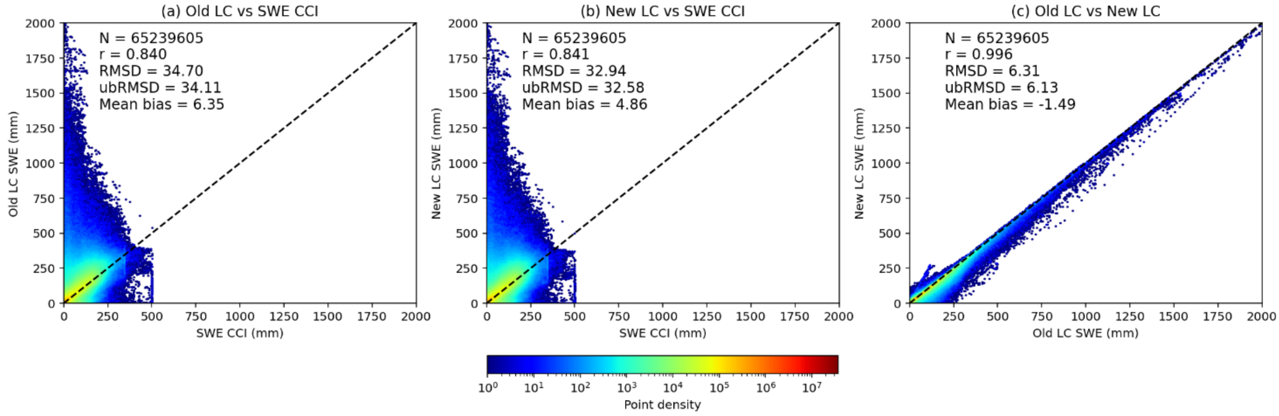

Figure R2.1 (New Fig. S1). *SWE scatterplots: (a) Old LC vs. CCI SWE , (b) New LC vs. CCI SWE, (c) Old LC vs. New LC, for the period 2010–2022..*

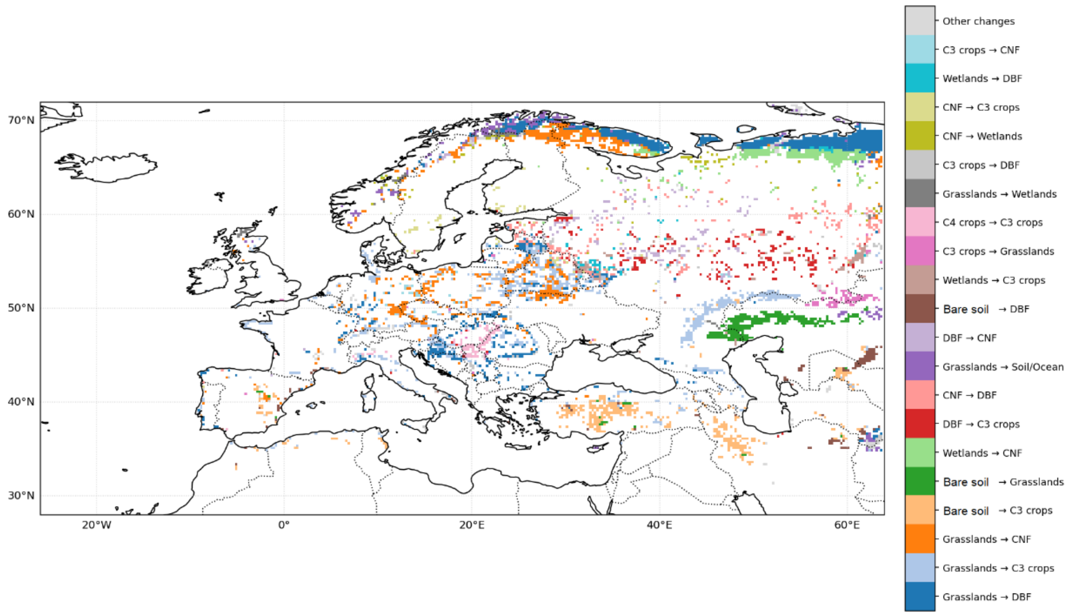

Figure R2.2 (new Fig. S4). *Spatial distribution of the top 20 most frequent dominant land cover transitions from Old LC to New LC over snow relevant grid cells where valid ESA CCI SWE data are available. All remaining transitions are grouped as "Other changes" Only transitions involving a change in dominant PFT are shown.*

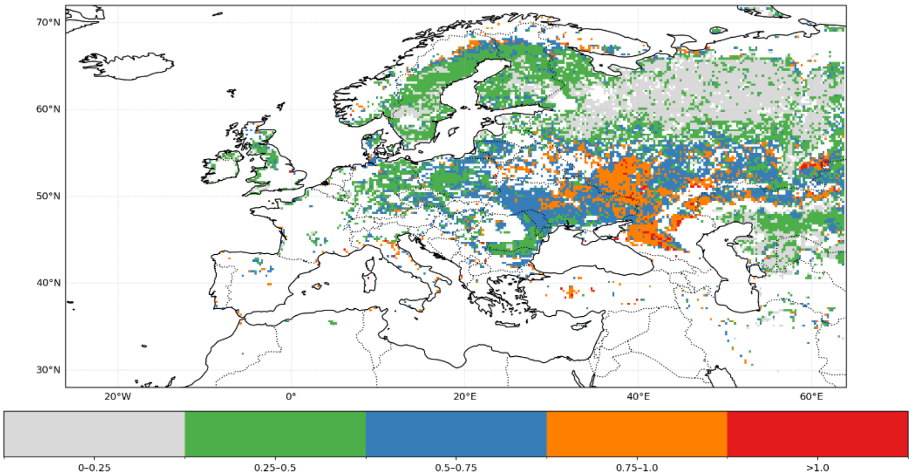

Figure R2.3 (New Fig. S5)*. Spatial distribution of grid cells with unchanged dominant PFT but modified sub-grid vegetation fractions. The colour classes represent intervals of the sum of absolute changes in PFT fractions. Values may exceed 1 due to the use of absolute difference s across multiple PFT fractions.*

Table R2.1 (New Table 2). Summary of the top 20 dominant land-cover transitions ranked by SWE RMSD changes. For each transition (Old LC to New LC), the domain-wide number and percentage of affected grid cells are reported, alongside the associated area-weighted reductions in SWE MB, RMSD and ubRMSD, in regions where snow is observed. Changes larger than 4 % are in bold. The remaining transitions are reported as "Other changes".

Transition	N° grid cells	% grid cells	Total contribution to Δbias (%)	Total contribution to ΔRMSD (%)	Total contribution to ΔubRMSD (%)
Grasslands -> DBF	932	4.8	**54.2**	**43.5**	**36.3**
Grasslands -> CNF	676	3.5	**10.3**	**11.8**	**12.0**
Grasslands -> C3 crops	822	4.2	**4.5**	**9.2**	**12.9**
Wetlands -> CNF	266	1.4	**6.8**	**6.5**	**6.6**
Grasslands -> Bare soil	188	1.0	3.8	**5.9**	**5.4**
Bare soil -> C3 crops	403	2.1	2.0	3.7	**5.3**
Grasslands -> Wetlands	86	0.4	2.3	3.5	3.4
CNF -> Wetlands	86	0.4	1.7	2.5	2.6
Bare soil -> Grasslands	278	1.4	1.0	1.6	2.0
C3 crops -> Grasslands	96	0.5	0.9	1.6	1.9
DBF -> C3 crops	237	1.2	1.3	1.5	1.6
Wetlands -> DBF	66	0.3	2.3	1.1	1.2
Wetlands -> C3 crops	109	0.6	0.4	1.0	1.3
DBF -> CNF	174	0.9	1.4	1.0	0.9
CNF -> C3 crops	75	0.4	0.5	0.6	0.7
Bare soil -> DBF	116	0.6	0.5	0.3	0.5
C3 crops -> DBF	86	0.4	0.6	0.3	0.4
C4 crops -> C3 crops	88	0.5	0.0	0.1	0.1
C3 crops -> CNF	34	0.2	0.1	0.1	0.1
CNF -> DBF	228	1.2	0.8	-0.6	-0.8
Other changes	147	2.8	**4.8**	**4.9**	**5.6**
Total	5193	28.8	100	100	100

- A new Section 5.1 was added to the manuscript.
- The Supplement was completed.

COMMENT 2.3. Which types of classification change experience the largest change in SWE? Were certain LC changes responsible for most of the change in MB?

RESPONSE 2.3:

The impact of LC change on SWE is strongly transition-dependent. A small number of transitions dominate the overall signal, exhibiting both a relatively large spatial extent and a strong local effect on SWE (see Table R2.1). The largest SWE changes are associated with transitions from grasslands to forest types, particularly the transitions from grasslands to DBF and CNF. As shown in Table R2.1, these transitions produce the strongest area-weighted reduction in SWE RMSD in regions where snow is observed (43.5% and 11.8%, respectively), and are also associated with substantial improvements in mean bias (54.2% and 10.3%, respectively), and ubRMSD (36.3% and 12%, respectively). Transitions from grasslands to C3 crops, wetlands to CNF, grasslands to bare soil and bare soil to C3 crops also make a significant contribution to the area-weighted reduction in SWE ubRMSD, at 12.9%, 6.6%, 5.4% and 5.3% respectively. These transitions are widespread across the landscape (see Fig. R2.2) and occur in regions where increased vegetation density and canopy effects significantly modify snow accumulation and melt processes. Notably, the sign of ΔBias (New LC − Old LC) is predominantly negative for these dominant transitions, indicating a systematic reduction in positive bias when using the New LC configuration. This suggests that the New LC configuration generally improves the representation of SWE magnitude in the affected regions compared to the Old LC configuration. Conversely, many other transitions exhibit only minor contributions, either because they affect a limited number of grid cells, or because their local impact on SWE is weak. These transitions can therefore be considered as having a negligible influence on the domain-scale bias. Overall, these results demonstrate that change in MB is largely controlled by a small number of meaningful vegetation transitions rather than being uniformly distributed across all LC changes.

- A new Section 5.1 was added to the manuscript.

COMMENT 2.4. From a qualitative perspective it appears that moving from bare soil (3) to deciduous broadleaf (9) had an impact on SWE. This should be discussed further.

RESPONSE 2.4:

We agree with the reviewer that this transition should be discussed in more detail. Transitions from bare soil to deciduous broadleaf forest (DBF) are present in the domain (see Fig. R2.2) and are associated with local changes in SWE that can be detected (see Table R2.1). These transitions are physically meaningful as they correspond to substantial modifications to surface properties, notably the introduction of vegetation where none was previously dominant. From a process-based perspective, the transition from bare soil to DBF affects several key mechanisms that control snow accumulation and ablation. Notably, the presence of a forest canopy results in the interception of snowfall, enhanced sublimation from the canopy and modifications to the surface energy balance through changes in albedo, longwave radiation and turbulent fluxes. These processes can lead to a reduction in snow accumulation on the ground and/or changes in melt dynamics, thereby impacting SWE. However, despite these clear local effects, these transitions only have a limited impact on the domain-wide SWE signal. As shown in Table R2.1, transitions involving bare soil (e.g. bare soil to DBF) account for only a small proportion of grid cells, contributing modestly to the total change in both RMSD and mean bias compared to more widespread transitions, such as grasslands to DBF or CNF. This suggests that, although bare soil to forest transitions can have a significant local impact on SWE, they do not dominate the large-scale response. Instead, overall SWE changes are primarily driven by extensive vegetation shifts that affect a larger portion of the snow-relevant domain. In the revised manuscript, we clarified this point by explicitly discussing the physical mechanisms associated with these transitions, and by contextualising their impact with respect to the dominant LC changes.

- A new Section 5.1 was added to the manuscript.

COMMENT 2.5. I am curious why areas that did not appear to have a change in LC (e.g. Sweden) seem to have more SWE in the ISBA run with the New LC compared to the Old LC. This should be examined.

RESPONSE 2.5:

In Sweden and Finland, it appears that the dominant PFT remained largely unchanged. In contrast, significant changes occurred in the other PFTs, as illustrated in Fig. R2.3.

COMMENT 2.6. Figure 3 and associated text: How did you address the absence of SWE estimates from CCI when producing Figure 3? Were the areas masked in CCI also removed from ISBA and ERA5?

RESPONSE 2.6:

Yes, this aspect was explicitly accounted for in our analysis. All of the comparisons presented in Fig. 3 of the main manuscript were performed using a consistent spatial mask derived from the ESA CCI SWE product. More specifically, we constructed a validity mask based on the presence of satellite data, retaining only grid cells for which at least one valid CCI SWE value was available over the study period. This mask was then systematically applied to all datasets, including the ISBA simulations (Old LC and New LC configurations) and ERA5. Consequently, the time series depicted in Fig. 3 of the main manuscript are computed over a shared subset of grid cells where the ESA CCI SWE product exists, thereby ensuring spatial consistency across all datasets. This avoids artificial biases that could arise from including model-only regions (i.e. areas without satellite coverage), particularly in regions with sparse or intermittent ESA CCI SWE data. In the revised manuscript, we clarified this point by explicitly stating that all model and reanalysis datasets are masked according to the availability of ESA CCI SWE data prior to computing spatial averages and statistics.

- Section 3 was revised.

COMMENT 2.7. L308-310: Limit is closer to 200mm than 100mm, although SWE exceeding 200mm is common in the CCI SWE product (see Fig. 4.10 [March] in Barella et al. 2024 and also Luojus et al. 2021). See Figure 4.15 in for spatial map of local biases.

RESPONSE 2.7:

We thank the reviewer for this important remark and for highlighting the occurrence of high SWE values in the ESA CCI product. As Luojus et al. (2021) discuss, the passive microwave signal used in SWE retrievals progressively saturates for deep snowpacks, corresponding to SWE values greater than 150 mm, which typically occur for snow depths of more than ~1 m. Under these conditions, the sensitivity of the signal to additional snow mass decreases, resulting in a systematic underestimation of SWE and increased uncertainty in the satellite product. To better assess this issue in our study, we analysed the distribution of SWE values in the ESA CCI dataset over the evaluation domain (see the histogram in Fig. R2.4). This analysis shows that the vast majority of SWE values are below 150 mm, with only a small fraction exceeding this threshold (approximately

3% of the total valid points in this case). Therefore, although high SWE values (>150–200 mm) do occur, they represent only a small part of the dataset and do not dominate the overall statistics. Since the passive microwave signal used in SWE retrievals saturates, the actual proportion of SWE values above 150 mm is higher. This threshold should not be interpreted as a strict upper limit, but rather as an approximate range beyond which the reliability of satellite-derived SWE decreases. We will revise the manuscript accordingly to clarify this.

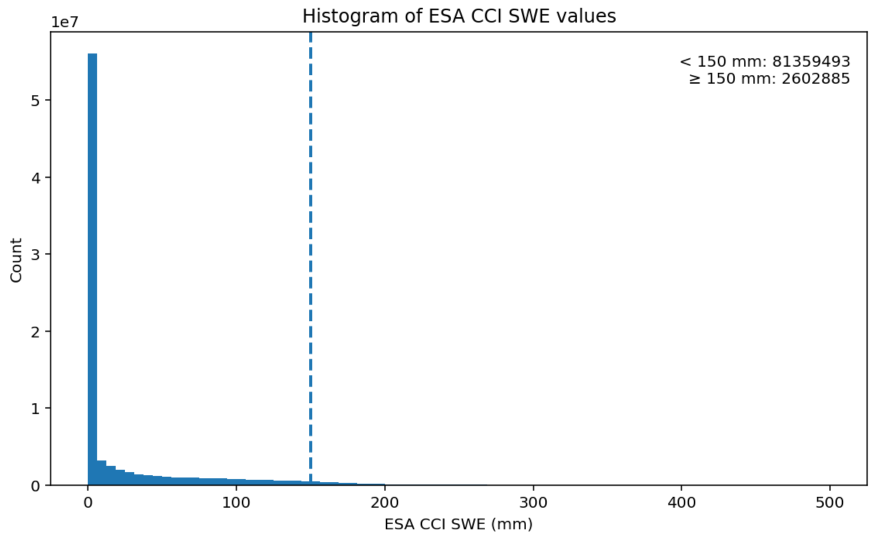

Figure R2.4 (New Fig. S7). *Statistical distribution of SWE values in the ESA CCI dataset over the evaluation domain.*

- Section 5.2 was revised accordingly.