Anonymous Referee #1

This study evaluated model parameterizations and configurations determining turbulent exchange and soil and vegetation and their influences on snowpack and surface energy fluxes using station-based eddy-covariance based energy flux data, snow depth, and soil temperature observations in two boreal and subarctic peatlands and forests in Finland. They found that a stability correction function is a key part to simulate sensible and latent heat fluxes over snow. Also, a realistic soil texture (soil organic carbon) parameterization led to improvement of the soil temperature simulations in peatlands.

The study presents insightful results describing model configuration in boreal peatlands and forests where are the most challenging regions to estimate snow energy balance budget. While the findings from this study should be tested by other snow model and LSMs to be widely generalized, as a case study, this study provides useful implications for choosing suitable turbulent flux parameterization and model structures to better estimate snow energy fluxes in soil-vegetation-snow interactions.

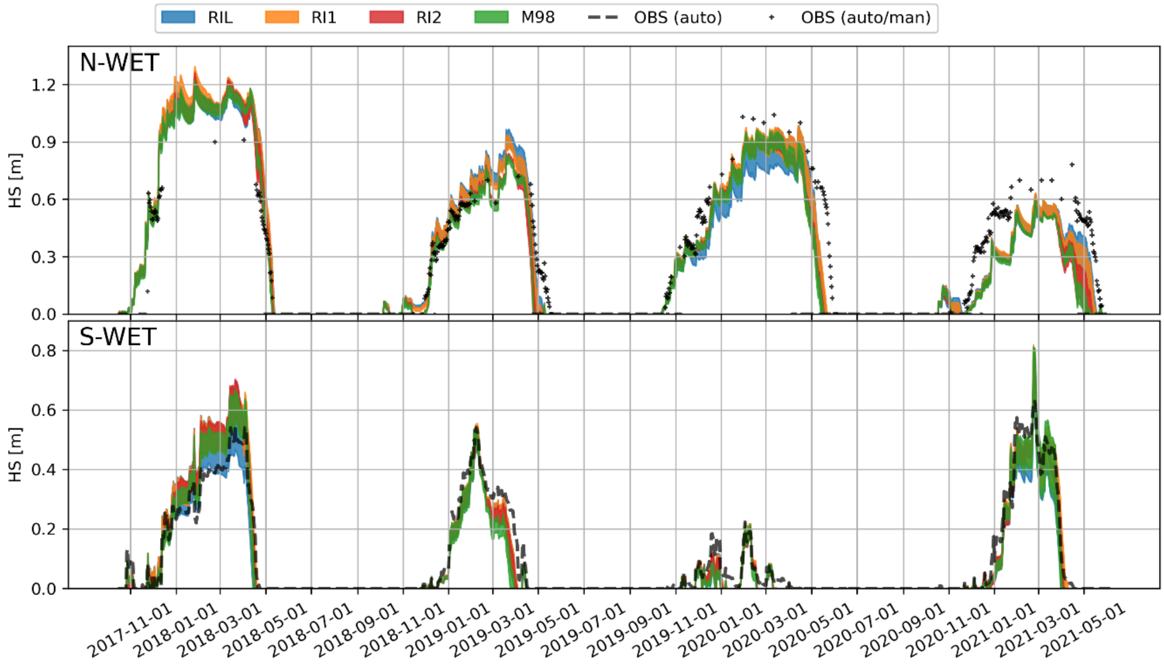
Even though I feel this manuscript is lengthy as a single manuscript, the paper is generally written well, and the presentation quality of the figures is great. In my opinion, however, concerns need to be addressed upon before publication is warranted. I have given a few suggestions below that may contribute to the improvement of the manuscript.

We would like to thank the Reviewer for the throughout assessment and positive comments regarding our study. We are happy that the Reviewer sees our results as insightful and the article as useful contribution to snow and LSM modelling community.

We would like to point out and apologize for two mistakes found in the data processing during an additional quality check motivated by the reviews. Thus, we would like to thank the reviewers for boosting our data quality control process.

First, we found 2h discrepancy in time zones between different data sources, notably between the main observations used for forcing and the data used for gap-filling. We have converted all the forcing and evaluation data into UTC time zone and ran new simulations. We will update the figures to include the corrected simulations in the revised manuscript. These changes do not have any major impact on the results of the study.

The second error was in the wind forcing reference height used for the S-WET site. The S-WET wind reference height had been assigned as in the contiguous S-FOR site (16.8 m), whereas in reality wind was measured at 3 m height. This means that simulated wind speed near the surface of the low-vegetated S-WET site was unrealistically low in our simulations. We have found that this reduces the sensitivity of turbulent fluxes (mainly sensible heat exchange) simulations to different turbulent exchange parameterizations, as the site experiences stable conditions less often than previously simulated. This correction does not change the results showing that M98 turbulence option is superior to classic stability correction function (RIL). However, this correction does change the snow depth simulations on S-WET. Figure 5. changes in a way that all the turbulent exchange parameterizations succeed in simulating the snow melt (see the revised Fig. 5 below). Results related to these changes will be corrected in the revised manuscript.



Revised Figure 5. Time series of snow depths simulated by ISBA-ESCROC. The 35 ensemble members are grouped by their turbulent flux parameterization, and the spread of each group is presented in colored ranges. Observed snow depths are presented in black dots and dashed lines.

Below we provide our detailed replies to general and detailed comments.

General comments

1. It is unclear to me about differences in MINERAL vs. SOC model runs to assess for assessing soil thermal regimes. Is the two model runs' difference soil texture only? In Table 2 of peatland sites, the authors assumed 100% SOC of soils. But in Figure 8, MINERAL was run as well. Does it mean the authors used "mineral soil" instead of 100% SOC with the same model? I don't know clearly how the ISBA parametrize soil texture but if "mineral soil" is continuous values (not categorized) within the model, it would be much more useful if you provide sensitivity of soil temperature and energy fluxes to different SOC percentages, e.g., 0, 20, 40, 60, & 100%).

This is a good point; the descriptions of peatland 'MINERAL' simulations were not addressed with sufficient detail and will be improved in the revised manuscript. These two simulations differ only in the soil texture. For the peatland 'MINERAL' simulation, we parameterized the soil texture (sand and clay fractions) identically as had been measured and used for the contiguous forest site (see Table 2.).

We are going to revise the Table 2. to be clear that the clay and sand fractions for peatlands are only effective below 1 m depth. The text at L340-L347 we are going to revise as:
"For a more detailed evaluation of two contrasting turbulent exchange options within ESCROC, we conducted deterministic (ISBA-FS) simulations with i) site parameters as shown in Table 2. and all default ESCROC parameterization as in Fig 2. in Lafaysse et al. (2017) (processes listed in Sect. 2.3.2, referred as RIL), and ii) site parameters as shown in Table 2. and all the default ESCROC parameterizations except the turbulent exchange option switched to M98 (referred as M98).

Moreover, we explore the influence of soil texture on the soil thermal regime and on snowpack dynamics. Hence, an additional deterministic simulation was conducted where the soil was characterized as mineral soil, identical as had been measured and used for the contiguous forest site, while turbulent exchange was set to M98 (referred as MINERAL). The MINERAL simulation was compared to the previously described M98 simulation, where the soil was characterized as fully organic until 1 m depth ~~1 m deep 100 % SOC~~ (Table 2)."

Rather than assessing uncertainties in quantifying the soil organic matter, the purpose of this sensitivity test was to show the consequences if the soil organic matter is neglected on boreal peatlands, which is the case in many applications of the SURFEX LSM. In addition, the content of organic matter in such boreal peatlands can be very close to 100%. For instance, Väliranta and Mathjissen (2021) measured mean loss on ignition value 96,6 % in the first meter of the peat profile at S-WET. Further, Muhic et al. (2023) also measured organic matter to be nearly 100% for the first 40 cm depth at a location just next to the N-WET peatland (see Muhic et al. 2023 Fig. 1b "FP2"). We are going to revise the manuscript by adding these citations to justify the use of 100% organic matter for model parametrization.

2. Even though the MINERAL vs. SOC comparison showed a more realistic soil texture (SOC 100%) provided a better performance, there were still unexplained portions of the modeled soil temperature as compared to observations. I would recommend including some discussions regarding the unexplained portions – how the model could additionally improve soil thermal regimes (e.g., soil moisture contents).

Indeed, the soil temperature simulations were still far from perfect and could be explained by some missing processes in the model. The missing processes were already discussed for snow cover, but we acknowledge discussion was not complete for potentially missing or not-well represented soil processes. In the revised manuscript, we will add a discussion about uncertainties in soil moisture content, which is especially important for the peatland as the moisture is expected to be controlled by water table dynamics, potentially affected by lateral flows not accounted for by this 1-dimensional LSM approach. Another source of uncertainty originates from the soil-vegetation composite approach where the vegetation layer is not explicitly represented above the soil. We will comment on this in the discussion of the revised manuscript as well.

3. I found "not shown" in several places. I understand the authors may want to keep the manuscript's length readable. However, "not shown" may not be helpful for potential readers who may want to keenly read this paper. I would recommend including any additional figures and tables that can support the findings from this study as Supporting Information (or Appendix).

Thanks for this comment, indeed we tried to keep the manuscript length readable by not showing additional figures. We have gone through our 'not shown' results and our plan is following:

- *Fig. 8: S-WET height of snow plot will be moved to Supplement. We will add previously 'not shown' N-WET height of snow to the Supplement plot as well.*
- *L464: Comparison of N-FOR simulated and observed SWE will be added to Supplement*
- *L493: "In the case of N-FOR, especially the summer energy fluxes were majorly improved by simply assigning the vegetation fraction to unity (full coverage, not shown)." This plot will also be added to Supplement*
- *L602: We believe 'not shown' is not necessary here and will be removed*

4. I wonder if there are any specific reasons focusing on 2018-19 winter. Was the winter a representative year in terms of snow climatology (e.g., moderate snow condition)? I would suggest providing a justification regarding that.

We chose this season as it contains the best coverage of energy flux data (least gaps) for all the sites, and it also represents well the typical snow conditions on the sites. We will clarify this in the revised manuscript.

Detailed comments

Figure 1. I think the current version of the figure does not extremely useful. Any detailed photos at the sites with positions of flux towers and soil temp & snow depth sensors may be needed. That would be helpful to better understand observations for those who are not familiar with those environments. Additionally, please include the site points on aerial photos with a scale bar. If the authors need more space, I recommend including them into Appendix.

Thank you, we have revised Fig 1. accordingly (see revised figure at the end), and it will be added to the revised manuscript. We are also going to add site photos in Supplement (see at the end).

L160 transpiration from the vegetation, Etr

Thanks for pointing out, this will be corrected.

L209 It would be good to include a citation (Vionnet et al., 2012) here

The citation will be included in the revised manuscript.

L222-223 My understanding is that at the bottom of the snowpack, Crocus is coupled to the soil components of the ISBA LSM. Please describe it explicitly.

What does "a mass and energy-conserving semi-implicit solution" mean? A detailed description of this would be helpful to understand.

As in many land surface models, the snow and soil modules in SURFEX are separated modules in which the thermal diffusion is first solved for snow and then for soil during the same time step. However, as the ground-snow heat flux is necessary at both steps, the thermal diffusion in snow is solved by using the first ground layer temperature from the previous time step and assuming it remains constant. The resulting snow-soil heat flux is then transmitted to the soil scheme to solve the heat diffusion in soil with a flux boundary condition, thus conserving energy in the coupling. Beyond this treatment of the soil-snow interface, the heat diffusion is solved in both snow and soil components with an Euler backward difference implicit numerical scheme (i.e. expressing snow/soil temperatures at the end of the time step in temperature evolution equations). We are going to summarize this in the revised manuscript at L221-223:

"The bottom of the snowpack and the uppermost soil layer are fully coupled with a mass and energy-conserving semi-implicit solution. The semi-implicit solution refers to a coupled system in which both components are solved separately with an implicit approach considering that the state of the second system remains constant during the solving of the first system."

L223 "Is there a specific equation of the heat conduction G related to the temperature gradient? I would suggest including this or citing a relevant reference.

The heat conduction flux is expressed using Fourier equation relating G to vertical temperature gradient and thermal conductivity as described in Eq. 4 of Decharme et al., 2011. We will add the reference to this Equation in the revised manuscript. The thermal conductivity and heat capacity are described using pedo-transfer functions (Noilhan and Mahfouf, 1996 and Peters-Lidard et al., 1998). These were already cited in the initial manuscript but will be introduced immediately after the introduction of heat conduction in the revised manuscript.

Did you use the heat conduction G as the same concept to the surface heat flux into the soil-vegetation composite (G; Eq 1)? If not, please use a different acronym.

G represents the total heat flux between two model components. In Eq. 1, it refers to the total heat flux between surface and atmosphere, while in Eq. 8, different subscripts indicate that it can represent the flux between the two first soil layers, two first snow layers, or between soil and snow. In the revised manuscript, all these terms will be more accurately defined and the subscript 0 will be added in Eq. 1 to better identify its meaning.

L291 What does "100% of SOC" mean? How much is this when it's converted as kg/m2? In Table 2, this should be expressed as kg/m2 for consistency.

We understand our definition of 100 % SOC was confusing. We meant that the soil is fully organic matter without any fraction of sand or clay. The 100% organic matter corresponds to 93,5 kg/m.^{2} We are going to revise our terminology and Table 2. accordingly.

Regarding the assumption, I would strongly suggest providing any justification of this assumption with sufficient references (e.g. why this is reasonable).

Please see our answer to this in the beginning of general comments. We are going to add justification in the revised manuscript.

L360 For model evaluation, have you considered daily-maximum and/or minimum values in addition to the daily-average values particularly for surface energy flux and soil temperature? Because of strong diurnal patterns of these variables, I think it would be useful.

Yes, we have considered maximum and minimum values as well but time series become quickly messy with too much information. The diurnal patterns are strong, and they are included in the scatterplot comparison of 6-hour averages of energy flux observations and simulations. See example Fig. 6, Fig. 7, Fig. 11 and Fig. 13.

Figure 7 please change the color of HS (obs) to black or others to make a difference from HS (mod).

The figure will be adapted as suggested

L425 Should "N-FOR" be corrected as "N-WET"? Figure 7 doesn't provide N-FOR time series

Thanks for pointing out this error which will be corrected in the revised manuscript

L426-426 I believe the authors used the ISBA-VS approach only on the peatland sites. But I'm curious how the results will be changed with ISBA-VS because this statement indicates that the uncertainty would be improved when the model considers fractional snow cover. Please consider

You probably mean that we only used ISBA-FS on peatlands. Using ISBA-VS in peatland sites would be less sensitive than on forest sites as the vegetation height is much lower, and therefore the snow fraction is quickly closer to 1 (i.e. identical to the ISBA-FS configuration). In any case, considering the interactions between snow and vegetation on peatlands would be likely improved using MEB as well *for low vegetation. However, as MEB has never been applied for snow-covered low vegetation, complementary developments and evaluations are required and this would be beyond the scope of our work. In-progress works in CNRM intend to generalize the applicability of MEB for these conditions. This point will be discussed in the revised manuscript.*

In Figure 7, during the melting period, there were different model errors in LSA between two sites (e.g., N-WET: underestimation but S-WET: overestimation). Please include some discussion regarding this.

In fact, the overestimation on S-WET is not during melt. This small overestimation is between the two main melt events, after a very light snowfall. We are going to add this in the revised manuscript.

L441 Please reword this sentence. It's unclear to me.

We will modify this to the revised manuscript as:

"This means soil temperature variations ~~became more rapidly~~are attenuated in M98 (including SOC) compared to MINERAL, and ~~the~~ this ~~effect~~ attenuation becomes increasingly important in deeper soil layers (Fig. 8)."

L443-447 The authors mentioned "On both sites" but a figure of N-WET wasn't presented here. I would recommend including the N-WET results as well (even in Supplementary Info) for those who are interested in. I'm not a big fan of "not shown".

We are going to add similar figure concerning N-WET soil temperatures in the Supplementary.

Did "SOC" simulation mean "M98" simulation (L345-347)? Because the acronym SOC also means "soil organic carbon" itself, it's a little bit confusing. Please use it for either one. Also clarify what the two simulations are fundamentally different in Sec 2.5. To my understanding, in MINERAL, the soil was characterized as "mineral" while SOC (or M98) as "100% SOC". Is this correct?

Indeed, SOC refers to simulation where soil is assigned as fully soil organic matter, and M98 turbulent flux option is used. Thus, SOC and M98 simulations are the same. As requested, we are going to change 'SOC' to M98 in the revised manuscript.

L456 I don't think Figure 9 shows the ISBA-VS model version heavily overestimates accumulation, except 2021 for N-FOR.

We will revise this by saying that the overestimation only occurs during 2021. This highlights the obscure results by ISBA-VS.

L456-458 Can you provide a figure comparing two runs with between Wsw = 5 and 0.2? What does it physically mean? In section 2.3.1, the author mentioned "The coefficient Wsw relates to vegetation characteristics" but I don't think it would be enough to physically understand the parameterization and the results here. Please provide more details.

Figure comparing Wsw = 5 and 0.2 is already provided in Appendix D, and is already at L459-L460: "The effect of snow cover fraction parameter wsw (see Sect. 2.3.1) for ISBA-VS snow depth simulations is detailed in Appendix D (Fig. D1)."

The idea of the Wsw parameter comes from the fact that the link between snow cover fraction and the vegetation characteristics are expected to be scale-dependent. Lower values of the parameter are expected at higher horizontal resolutions (LSM grid size). However as this parameterization has never been calibrated or evaluated against observed snow cover fractions, it can be seen from Table

D1 that this scale-dependence was not consistently applied among all applications, and that this parameter is sometimes adjusted on other criteria. This information was provided by developers of ISBA who have published the parameterization but without this level of detail, which may partly explain the lack of consistency among the different applications.

Reviewer 2 used "tweaking parameter" to qualify Wsw which is unfortunately common in numerical models. Our idea here is to emphasize these inconsistencies and to illustrate their impact on simulation results, see also our response to Reviewer 2. The discussion on that topic will be improved in the revised manuscript.

L462-463 How did the authors know that the results are due to overestimated compaction and uncertainties in snowmelt? For example, have you compared snow density (as a proxy of compaction) with observations? If so, why did the errors of compaction and melt processes occur in terms of model physics - particularly snow energy processes?

We also compared the simulated and observed snow water equivalent (SWE) on N-FOR but did not include the comparison in the manuscript as the SWE observations were not continuous (and not to further lengthen the manuscript). This SWE comparison will be included in the Supplement of the revised manuscript (see also Fig. 1 below), and it shows that the SWE is rather well simulated during the accumulation period. However, as the SWE does not fully explain the errors in HS simulations (underestimated peak HS and too early decrease in peak HS), the errors in snow simulations must also come from compaction. Uncertainties in compaction routine of Crocus are known as compaction due to metamorphism is not represented (Vionnet et al. 2012, Lafaysse et al 2017). Further, due to the interaction between compaction and surface energy budget, both errors may be linked. First, incorrectly simulated surface energy balance influences the compaction (e.g. warmer snow or higher liquid water content increases compaction). Secondly, errors in densification of snow have an impact on the surface energy budget as thermal conductivity of snow depends on density.

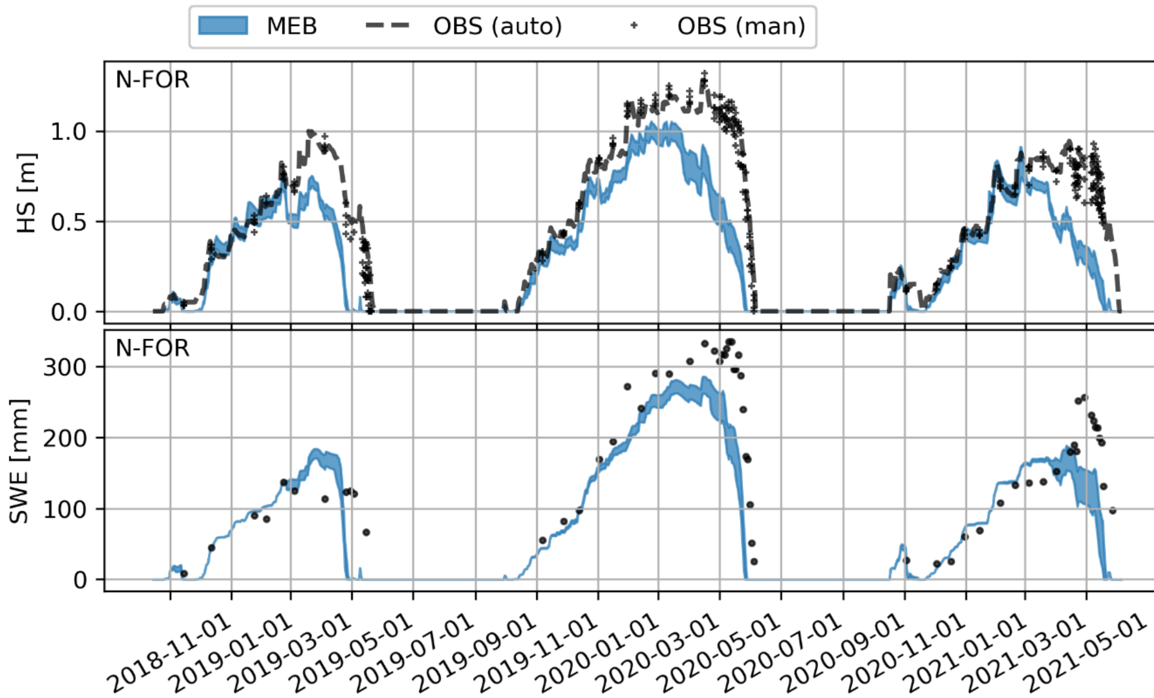


Figure 1. MEB simulations against observed height of snow (HS) and snow water equivalent (SWE) on N-FOR site.

L464 Again, please include your relevant result in Supporting Info instead of "not shown". I think the authors would be able to discuss more in this section (3.3.2) because there are lots of interesting patterns such as the cold biases in fall. Can you add air temperature in this figure for comparison purposes? I'm guessing that the modeled soil temp seems to be much more sensitive to air temperature. Also, I would suggest adding discussion about the differences in soil temperature between ISBA-FS and MEB.

Please see our earlier answer concerning all 'not shown' notions. We agree that air temperature could be interesting addition to the plot, but it would also require significant increase of the scale, making the soil temperature dynamics less visible. We are going to consider this for the revised manuscript. We are going to add the suggested discussion in the revised manuscript,

L489-491 So how would the roughness length and turbulent exchange coefficient (CH) be used differently for soil and vegetation to minimize the overestimations of soil evaporation and sublimation?

This is a good point and part of the motivation for developing MEB, so that the vegetation and soil can have separate roughness lengths, turbulent exchange coefficients, and thus, different magnitudes of turbulent fluxes. So, the answer is 'by using an explicit vegetation model' (see more details in Boone et al. 2017)

Figure 12, for N-FOR, were the LSA observations missing or larger than 0.7? Please justify. To me, it is questionable why the LSAs from models were constant throughout the season, likely regardless of the existence of snowpack, which don't seem to make sense. Please discuss the potential causes.

The gap in the N-FOR LSA observations is due to the polar night. The simulated LSAs for the forest are dominated by the albedo of vegetation that was assigned as constant (see Table 2.). The influence of evolving albedo of snow is thus not visible for the LSA.

L498 "neglect" -> "do not estimate"?

We will modify this in the revised manuscript

L547-549 I don't think this statement is sufficiently supported by this study because the snow depth patterns from the models highly depend on year-by-year (Figures 5 & 9). "Wind erosion is higher in peatlands" is generally correct, but have you seen the wind data in the sites to see any wind effect?

Thanks for pointing this out. We agree that the extent of wind erosion cannot be derived from the data, and the argument was too speculative. We will revise the text.

Table 3 "Weakly stable" sounds somewhat weird. Please consider an alternative.

The terminology used here is common for characterizing the stability regimes for atmospheric boundary layer flows (see e.g. Grachev et al. (2005, Boundary Layer Meteorology 116, p.201–235). We recognize that this may differ among research disciplines, however.

L584-587 Does Crocus itself provide soil temperature simulation? I believe it only estimates snow and is coupled with ISBA to estimate soil parts. Thus it would be better to state something like "ISBA coupled with Crocus"

You are right, the soil heat budget is solved in ISBA. We will revise this sentence to avoid confusion.

L597 I don't think the authors explicitly showed soil evaporation from the results. Please rephrase it.

Thanks for pointing out this inconsistency. Although we did not explicitly show soil evaporation in the results, the finding that the N-FOR latent heat simulations were improved when the vegetation fraction was increased (0.95 to 1.0, i.e all ET becomes controlled by stomatal), suggests that the soil evaporation plays a too big role. The sentence will be revised accordingly.

"We found ISBA-VS to drastically overestimate the LE, likely because of too high soil evaporation simulations due to its conceptualization of vegetation and snow cover fraction~~, resulting in too high LE~~."

L717-718 Please reword

We will modify this in the revised manuscript as:

"Our evaluation with the ~~ESCROC~~ ensemble ~~Crocus~~ snowpack model ~~framework~~(ESCROC) ~~gives confidence~~ ensures that uncertainties in snow processes (not evaluated in this study) do not affect the robustness of our main conclusions ~~summarized below~~."

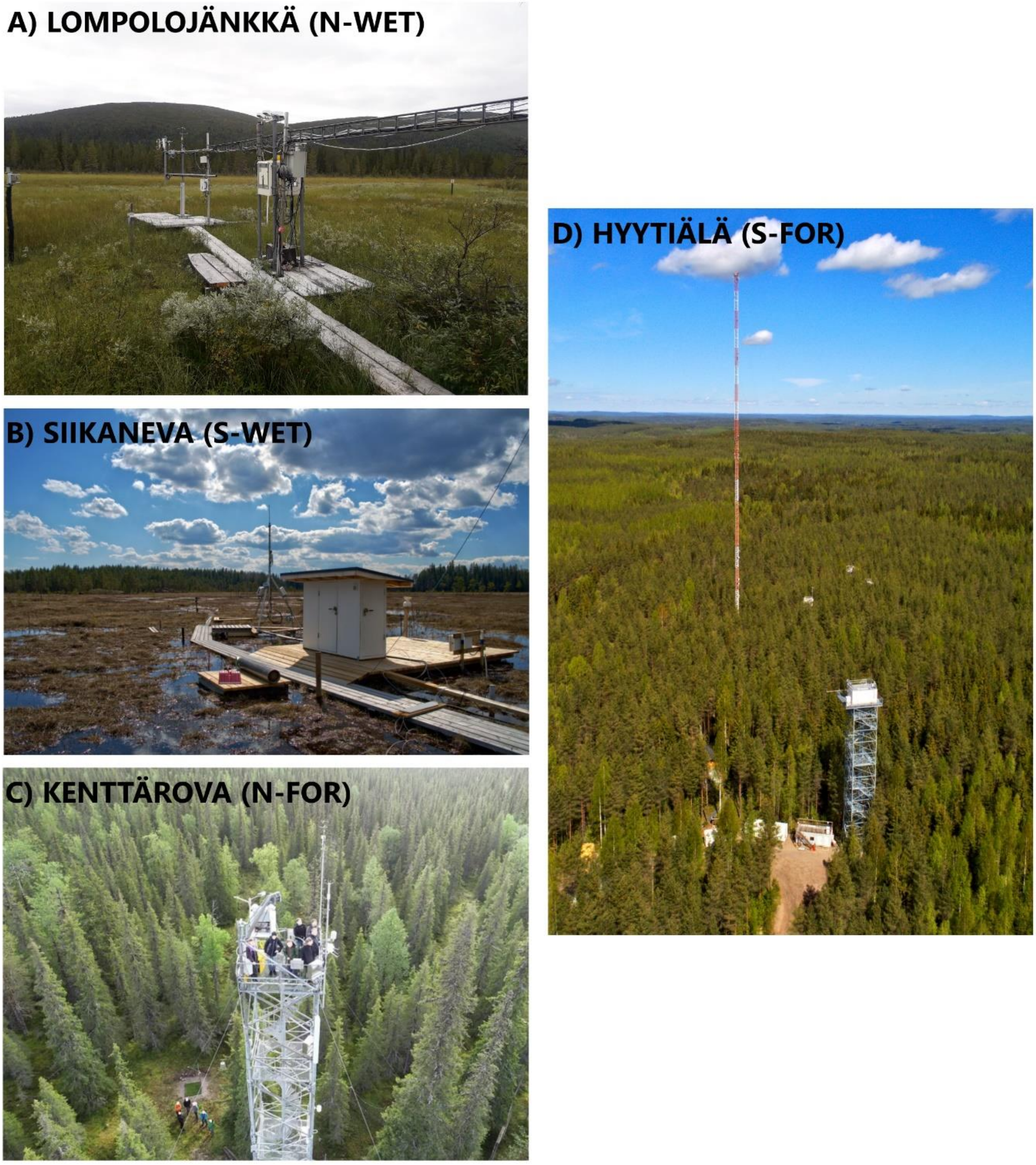


Figure 2. Study site pictures (Lompolojänkkä (Pertti Ala-Aho), Siikaneva (Alekseychik et al. 2022), Kenttärova (Bastian Steinhoff-Knopp), Hyytiälä (Kolari et al. 2022)).

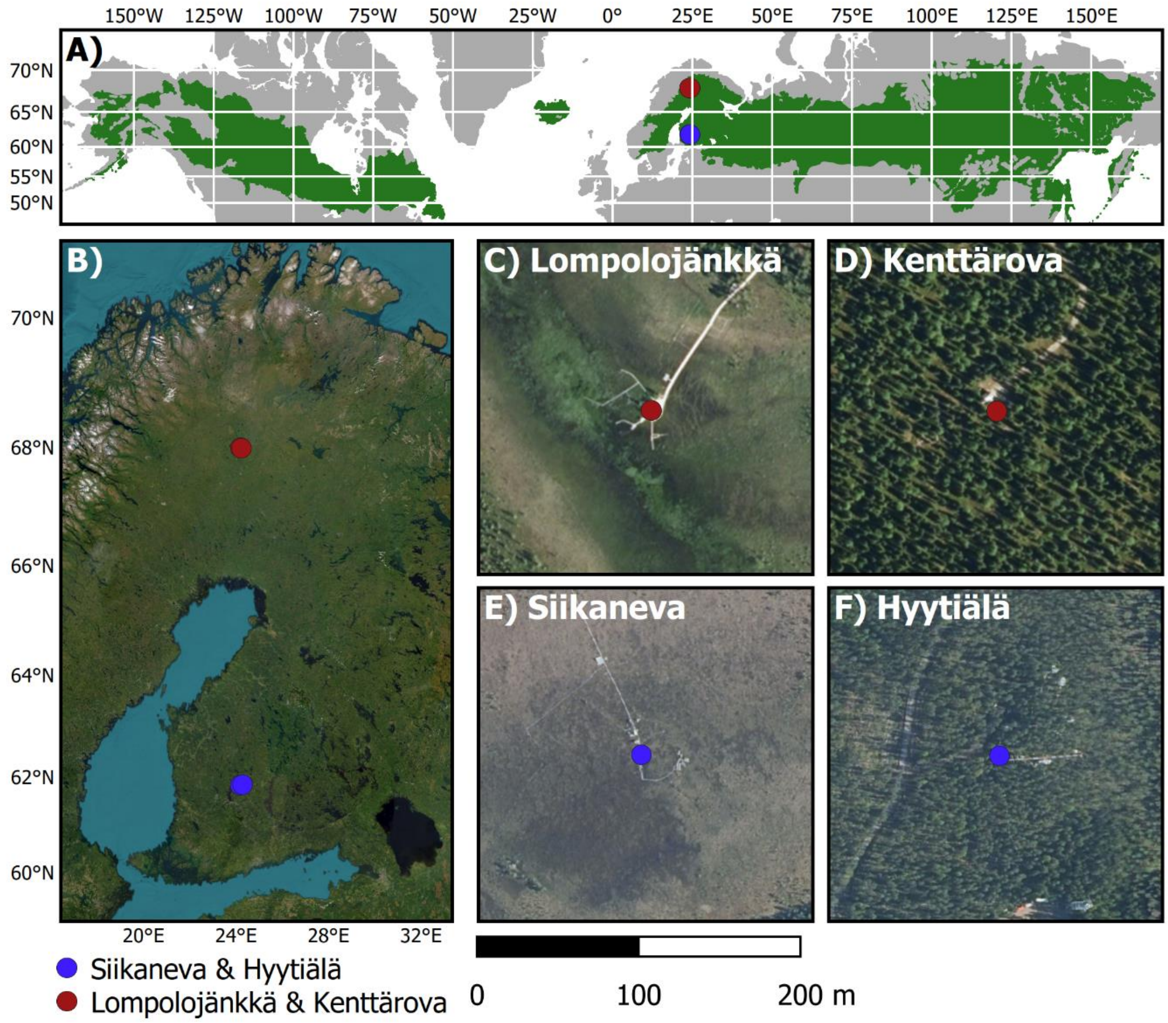


Figure 3. A) Study area locations inside the boreal land biome (green area, Olson et al., 2001), B) study sites locations in Finland (Source: Esri) and C-F) aerial images of each site (NLSF, 2020).

References:

Grachev, Andrey A., Christopher W. Fairall, P. Ola G. Persson, Edgar L. Andreas, and Peter S. Guest. "Stable boundary-layer scaling regimes: The SHEBA data." Boundary-Layer Meteorology 116 (2005): 201-235.

Boone, A., Samuelsson, P., Gollvik, S., Napoly, A., Jarlan, L., Brun, E., & Decharme, B. (2017). The interactions between soil-biosphere-atmosphere land surface model with a multi-energy balance (ISBA-MEB) option in SURFEXv8-Part 1: Model description. Geoscientific Model Development, 10(2), 843–872. https://doi.org/10.5194/gmd-10-843-2017

Väliranta, Minna; Mathijssen, Paul J H (2021): Geochemistry of Siikaneva peat core from Finland. PANGAEA, https://doi.org/10.1594/PANGAEA.927689

Muhic, F., Ala-Aho, P., Noor, K., Welker, J. M., Klöve, B., & Marttila, H. (2023). Flushing or mixing? Stable water isotopes reveal differences in arctic forest and peatland soil water seasonality. Hydrological Processes, 37(1), e14811. https://doi.org/10.1002/hyp.14811

Noilhan, J., & Mahfouf, J. F. (1996). The ISBA land surface parameterisation scheme. Global and Planetary Change, 13(1–4), 145–159. https://doi.org/10.1016/0921-8181(95)00043-7

Decharme, B., Boone, A., Delire, C., & Noilhan, J. (2011). Local evaluation of the Interaction between Soil Biosphere Atmosphere soil multilayer diffusion scheme using four pedotransfer functions. Journal of Geophysical Research Atmospheres, 116(20), 1–29. https://doi.org/10.1029/2011JD016002

Peters-Lidard, C. D., Blackburn, E., Liang, X., & Wood, E. F. (1998). The effect of soil thermal conductivity parameterization on surface energy fluxes and temperatures. Journal of the Atmospheric Sciences, 55(7), 1209–1224. https://doi.org/10.1175/1520-0469(1998)055<1209:TEOSTC>2.0.CO;2

Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., Powell, G. V. N., Underwood, E. C., D'Amico, J. A., Itoua, I., Strand, H. E., Morrison, J. C., Loucks, C. J., Allnutt, T. F., Ricketts, T. H., Kura, Y., Lamoreux, J. F., Wettengel, W. W., Hedao, P., & Kassem, K. R. (2001). Terrestrial ecoregions of the world: A new map of life on Earth. BioScience, 51(11), 933–938. https://doi.org/10.1641/0006-3568(2001)051[0933:TEOTWA]2.0.CO;2

NLSF. (2020). National Land Survey of Finland Topographic Database. Available at: Http://Www.Maanmittauslaitos.Fi/En/e-Services/Open-Data-File-Download-Service.