Response to Reviewer 1

I thank the reviewer for his / her valuable comments on the manuscript. My response to the comments and the changes I plan to make in the revised manuscript are detailed below. For clarity, the comments are in blue font, while my response is in black. In some cases, I have included text planned to appear in the revised manuscript in red font.

The author discusses the mechanisms behind interannual variability in snow depth / SWE in Northern Europe, and how these mechanisms differ from that of the projected long-term climate change. Analysis is performed on the ERA5-land data set and regional climate models from EURO-CORDEX, by an earlier published method decomposing the contributions from change in total precipitation, snowfall fraction and snow-on-ground fraction (melting). The paper is well written, easy to follow and fit well in The Cryosphere. The work is relevant for everyone that wants to understand the mechanisms behind changes inC1snow on ground in more detail and explains why "warm winters" of today's climate can't be used as an analogy to future projected climate change with respect to snow cover. In general I like the paper very much, but have some comments and questions which hopefully will improve and clarify certain aspects

General comments

The scientific problem investigated is in itself interesting and line 59-60 is a start to put the work in a broader context and meaning. However, in the introduction I miss some further elaboration on why this work is important for a wider audience. Related to this, in the conclusions, is it possible to say something on the consequences of the findings in the study. For example (my understanding) does it mean that the accuracy of the snow schemes (i.e. melting processes) is important to get correct to simulate future changes in snow amounts in northern Europe?

This is a good point. I plan to add the following paragraph towards the end of the Introduction:

The significance of this research in a wider perspective is twofold. First, a better understanding of the processes involved in the interannual variability and long-term trends of snow conditions is valuable for model developers, helping to focus the development work towards the most important processes. For example, the findings in this paper suggest that, in areas with relative mild winters like southern Finland, the description of snowmelt is important for the simulation of both the interannual variability and future trends of snow amount. Second, the current results bear an important message for climate impact researchers and the general audience, by showing why the snow conditions in individual mild winters are not a perfect analogy for what to expect in the future.

The study refers to some earlier works, but not very many.

However a google search shows many results on scientific papers on snow cover in Europe and also with examples of using data from, e.g. Sodankyla. This reviewer doesn't know the details of these papers, but imagine that at least some of them might be relevant to mention in the context of this work. For example the recent Cryosphere study looks relevant (?)

Essery, R., Kim, H., Wang, L., Bartlett, P., Boone, A., Brutel-Vuilmet, C., Burke, E.,Cuntz, M., Decharme, B., Dutra, E., Fang, X., Gusev, Y., Hagemann, S., Haverd,V., Kontu, A., Krinner, G., Lafaysse, M., Lejeune, Y., Marke, T., Marks, D., Marty,C., Menard, C. B., Nasonova, O., Nitta, T., Pomeroy, J., Schädler, G., Semenov, V.,Smirnova, T., Swenson, S., Turkov, D., Wever, N., and Yuan, H.: Snow cover duration trends observed at sites and predicted by multiple models, The Cryosphere, 14,4687–4698, https://doi.org/10.5194/tc-14-4687-2020, 2020.Even if it won't change the results from this study it would be nice to tie the presentwork more together with a broader part of earlier work.

Thanks for this reference. The mentioned study nicely exemplifies how off-line comparison of land surface models could be used to study the uncertainty in such processes (e.g., snow melt) for which direct observations are not available. I plan to include in the concluding section of the revised manuscript.

Considering the comments of all three reviewers and my own literature search, I also plan to add 10-15 other new references in the revised manuscript.

The main work in this study relies heavily on the ERA5-Land data set. Some verification of ERA5-Land is also included at Sodankyla and Helsinki. The results of theverification in itself are encouraging. However, only comparing ERA5-Land with observations at two locations when discussion the results also in other terrain types, e.g. Scandinavian mountains are a bit scarce. I can imagine that the errors might be larger at other locations. If other studies of the quality of ERA5-land for northern Europe exist they could be referred to.

ERA5-Land is still a new data set, and therefore few verification studies are available. I have found a couple of examples that I may mention in the revised manuscript, but I am not aware of anything focussing on northern Europe.

Also, it is true that the verification provided in the manuscript is far from exhaustive. This will be clearly acknowledged in the revised manuscript, where I will also discuss some further aspects that could (or at least ideally should) be addressed in further work, approximately as follows (after Fig. 3 in Section 2):

The comparison presented in Fig. 3 is far from exhaustive. More insight could be gained, for example, by extending the evaluation to the daily time scale, but this is out of the focus of the present study. Another unverified aspect is the ability of ERA5-Land to distinguish between solid and liquid precipitation in near-zero temperatures. This is important because, in principle, a good simulation of snow amount might still hide compensating errors in snowfall and snowmelt. Unfortunately, there is no ground truth to compare with, since precipitation measurements in Finland only record the total precipitation. Empirical estimates for the dependence of the snowfall/rainfall probability on near-surface temperature and humidity have been derived based on synoptic observations (e.g., Auer, 1974; Koistinen et al., 2004), but the conversion to total daily snowfall or rainfall fractions is nontrivial because precipitation intensity, temperature and humidity all vary on sub-daily time scales.

I also miss some verification of the precipitation phase in ERA5-Land, it should be possible to compare from observations, either by manual observations (if available) or by using some temperature thresholds based on the observations.

I agree that the ability of ERA5-Land to represent the phase of precipitation is one of the main uncertainties in the study. The problem is that verification of this is far from trivial. The visually or optically determined phase of precipitation is encoded in the present-weather code in SYNOP observations, but in practice these data are not easily available, would require substantial work to process, the processing would require subjective choices (e.g. how to deal with sleet?), and the conversion to precipitation totals would require information on precipitation intensity on sub-daily time scales. Approximate temperature-based methods (as mentioned in the previous red paragraph) might be slightly more straightforward to use, but still represent an approximation and would require data on temperature and precipitation intensity on sub-daily time scales. Considering these issues, I feel it is better to leave this as a topic for further research. However, I will acknowledge this issue in the revised manuscript (as indicated in the paragraph in red above).

Finally for the verification part, is it possible and meaningful to estimate any of the decomposed contributions to snow variability from observations and compare them with ERA5-Land?

This would certainly be useful, but is far from trivial in practice. The problem is the lack of direct observations for the snowfall fraction of total precipitation (as discussed above), together with the fact that measurements of SWE are scattered in space and available at irregular time intervals.

While the presented analysis relies on mean temperatures, how will the local temperature variability impact, e.g. in some areas there might be large variability with several “above zero degree” periods while in other areas it will be less variability and may beconstant temperatures under zero with the same mean temperature? I’m not suggesting any new analysis, but I’m curious about if you have any opinions on this subject?

Can this partly explain why the correlation of Fig2a doesn't follow the mean temperatures?

Differences in temperature variability may well play a role, but it is difficult to separate this from other factors. Among them, differences in total winter precipitation (i.e. much more precipitation falling in western Norway than elsewhere in the study area) appear to play an important role. To demonstrate this, and also to answer your last comment on what is meant by "colder areas", I plan to add the following new figure to the revised manuscript:

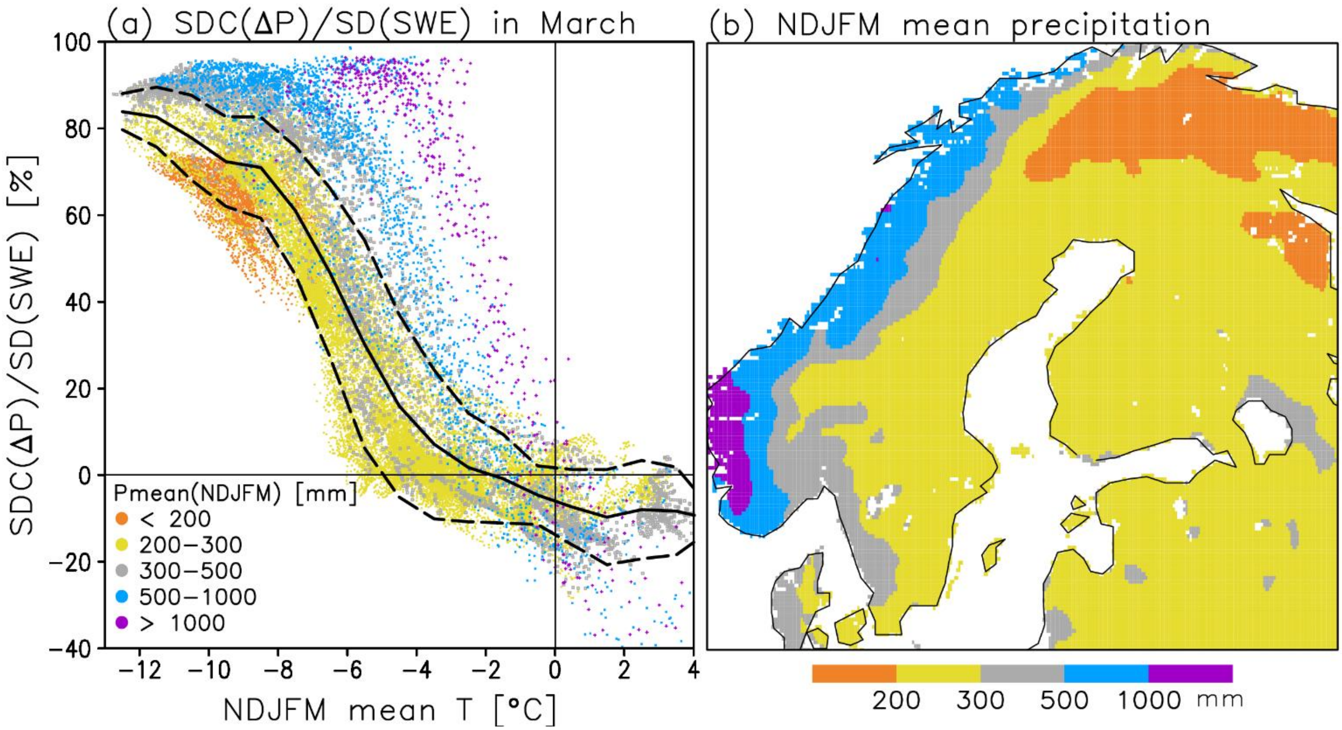

Figure 8. (a) The relative contribution of precipitation anomalies to the standard deviation of SWE in March as a function of the NDJFM mean temperature in 1981/82 – 2019/20. Each dot represents a single 0.1° × 0.1° grid cell, coloured according to the mean NDJFM precipitation shown in (b). The solid line in (a) indicates the mean values for 1°C temperature bins, and the two dashed lines the mean ± one standard deviation.

together with the text in the next two paragraphs (the first one devoted for quantifying the temperature dependence and the second discussing the role of total precipitation):

As a further illustration, the relative contribution of precipitation variability on SWE variability in March (row 3, column 3 in Fig. 7) is plotted as a function of the climatological NDJFM mean temperature in Fig. 8a. On the average, this contribution exceeds 80% where $T_{NDJFM} < -11$°C, is close to 50% where $T_{NDJFM} \approx -7$°C, and decreases to zero at $T_{NDJFM} \approx -2$°C. Despite the non-linearity of the relationship, there is a strong negative spatial correlation (*r* = -0.85) between the two variables in Fig. 8a. Conversely, the relative contributions of snowfall fraction variability ($sdc(\Delta F)/sdc(SWE)$) and snow-on-ground fraction variability ($sdc(\Delta G)/sd(SWE)$) are positively correlated with the NDJFM mean temperature (r = 0.65 and 0.83, respectively).

Nevertheless, the dynamics of interannual SWE variability is not solely controlled by the winter mean temperature. For the same NDJFM mean temperature, and excluding the mildest areas, $sdc(\Delta P)/sd(SWE)$ tends to increase with increasing NDJFM mean precipitation (see the colour coding in Fig. 8a). In particular, the SWE variability in western Norway, where more precipitation falls than elsewhere in Northern Europe (Fig. 8b), is more strongly affected by precipitation variability than expected from the winter mean temperature alone. On one hand, the larger mean precipitation is associated with larger absolute precipitation variability. On the other hand, larger amounts of snowfall reduce the variability in the snow-on-ground fraction, because a larger amount of snowmelt is needed for a unit change in the latter.

Specific/minor comments;

Line 26: Since the example is one single winter (2019/20) it is not an example of largeinterannual variations in my view. Is it better to write something like: “An example of a particular anomalous winter was the winter 2019/20 . . . .”?

Will be changed approximately as suggested.

Line 28-29: I suggest to add some details about the regional contrasts in the text, notonly refer to Fig 1c. That would improve the flow in the reading.

I will add something like this: Record-breaking positive anomalies of 3-5°C in the November-to-March mean temperature extended from southern Sweden to southern and central Finland, the Baltic States and western Russia, whereas the precipitation surplus was unusually large especially in Finland.

Line 30-31: Add “(indicated by stippling in Fig 1c)” at the end of the sentence? I used some time to wonder how you knew that this was record low.

Will be done.

C3Line 42: Can the word “interannual” be skipped in this sentence?

Yes, I will skip it.

Line 68-69: Since SWE and snow depth are not equal they are not necessarily straightforward to compare, e.g. a change in snow depth can also come from changes in snow density. Is this important or discussed anywhere in the references?

I will add a brief note ("Note, though, that snow depth is affected by snow density as well as SWE") on this in the text.

Line 75: In my first read I wondered what the exact meaning of "the fraction of accumulated snowfall that remains on the ground" was. It is explained better later, but can something to clarify what it is be added already here (e.g. "the partition not melting")?

I will add a brief explanation as: the fraction of accumulated snowfall that has not yet melted and thus remains on ground (snow-on-ground fraction)

Line 102: Without having detailed knowledge about the local terrain/weather around Helsinki and Sodankyla I wonder if at least the coast line near Helsinki introduces somegradients in temperature? But as it is commented, since both Helsinki and Sodankylaobservations have a long observation track and distribute observations via GTS theyare probably assimilated in ERA5. I think that in combination with the fact that you com-pare monthly mean and not hourly values ensure the very high correspondence with observations? I guess the correspondence with observations are reduced somewhat ifhourly values are compared.

It is indeed possible, and even likely, that the specifics of the local station environments explain some of the apparent biases in ERA5-Land. It also seems likely that the correspondence on daily or hourly time scales would be worse than that for monthly or winter mean values. I will mention these issues briefly when discussing Figure 3 in the revised manuscript, but only briefly as these details are somewhat out of the focus of the study.

Line 103-106: I agree with the reasoning and think that it is actually quite difficult to say if the true bias of ERA5 precipitation is positive or negative. Can you estimate how large, in percent, the apparent bias is? That can make it easier to judge (if you also know the gauge equipment characteristics and wind climate). I don't think it is necessary to be very quantitative, but it could strengthen your qualitative statement.

Thanks for this question. A closer look at Taskinen and Söderholm (2016, full reference in the original manuscript) allows the following addition at this point: "The absolute mean values in the reanalysis exceed the station measurements by 12% in Helsinki and 18% in Sodankylä, … In fact, the difference between ERA5-Land and the station observations agrees well with

Taskinen and Söderholm (2016), who estimate the average December-to-March precipitation in Finland and its cross-boundary watersheds in 1982-2011 to have been 17.5% larger than measured".

Figure 3: I think you can argue that the temperature bias is smaller during precipitation events. In particular the positive bias in Sodankyla probably arises from stably stratified situations with little precipitation.

This is plausible. However, to keep the focus of the paper, I probably won't comment on this in the revised manuscript – in particular as the cold winter climate in Sodankylä means that snow amount is more strongly controlled by precipitation than temperature.

Equation 2: It would be nice with some more details on how the different terms arecalculated from ERA5-Land and the RCMs. For example, for the snowfall, do you use snowfall from ERA5-land (?) or do you use some temperature thresholds on theC4 total precipitation to decide precipitation phase?

To clarify, this addition will be included to the paragraph following Eq. (1): "All the variables required in Eq. (1) (i.e., the total precipitation, snowfall and SWE) are directly available for both ERA5-Land and the EURO-CORDEX simulations."

Are snowfall available from all RCMs? Even more, I'm curious on how you calculate the G-term in practise. Can you give some more details?

Yes, snowfall is available for all the RCMs included in this study (see above). The snow-on-ground fraction *G* is obtained by dividing SWE by the accumulated (i.e. time integrated) snowfall. I will add a brief explanation on this below Eq. (1) in the revised manuscript.

Line 150-151: Isn't "1" applied on single winters and then the average is taken overall these winters? Could the word "decomposed" be added in front of "values for anindividual winter"?

Sorry for being unclear with the notations. "1" and "2" here referred to the subscripts introduced just above Eq. (1). The planned clarification for this is as follows:

In this study, the decomposition (2) is applied in two different ways:

1. When studying interannual variations in SWE, X_1 as defined above Eq. (2) represent the mean values for a 39-winter period (1981/82 to 2019/20, 2020/21 to 2058/59 or 2059/60 to 2097/98) and X_2 the values for an individual winter.

2. When studying long-term changes in SWE, X_1 represent the mean values for winters 1981/82 to 2019/20, and X_2 those for either 2020/21 to 2058/59 or 2059/60 to 2097/98.

Line 152-153: Again isn't "1" applied on single winters and then averaged? Or do I misunderstand?

See the clarification above. Your misunderstanding was entirely my fault!

Equation 3 & 4: I don't understand these. I think they need some more explanations.That would help me to understand e.g. Line 206-207 also.Figure 7: What does the "C" in "SDC" stand for?

SDC stands for "standard deviation contribution". I will add a note on this. Also, to make the statistics easier to follow, I will provide a longer derivation in the revised manuscript, as follows:

Multiplying Eq. (2) with ΔSWE and averaging over a 39-winter period, the interannual variance of SWE can be decomposed to the contributions of the four right-hand-side (rhs) terms in Eq. (2) as

$$var(SWE) = \langle \Delta SWE^2 \rangle = \langle \Delta SWE \sum_{i=1}^{4} \Delta SWE_i \rangle = \sum_{i=1}^{4} cov(\Delta SWE_i, SWE) \tag{3}$$

where the angle brackets indicate a time mean, *var* is variance and *cov* covariance. Similarly, the standard deviation (s) of SWE is decomposed as

$$s(SWE) = \frac{var(SWE)}{s(SWE)} = \sum_{i=1}^{4} \frac{cov(\Delta SWE_i, SWE)}{s(SWE)} = \sum_{i=1}^{4} \frac{cov(\Delta SWE_i, SWE)}{s(SWE)s(\Delta SWE_i)} s(\Delta SWE_i) \tag{4}$$

which can be rewritten using the definition of correlation (r) as

$$s(SWE) = \sum_{i=1}^{4} sdc_i = \sum_{i=1}^{4} r(\Delta SWE_i, SWE) s(\Delta SWE_i) \tag{5}$$

where the sdc_i:s refer to the *standard deviation contributions* of the four rhs terms in Eq. (2).

Line 261: add "change" after "SWE"?

Suggestion accepted.

Figure 10. The discussion of Fig 10 is not clear to me. I don't fully understand what the message is. Can this be made clearer?

The main point in Fig. 10 (to be Fig. 11 in the revised manuscript) is the difference between the c- and d-panels, which shows that, for each 1°C of warming, the projected long-term increase in precipitation is different, and generally smaller, than the precipitation anomalies accompanying a 1°C interannual winter temperature anomaly in ERA5-Land. Here is the planned revision, also taking into account the comments of the other reviewers:

In apparent conflict with the simulated future decrease in SWE nearly everywhere in northern Europe, Fig. 2a showed a positive interannual correlation between March mean SWE and NDJFM mean temperature over the Scandinavian mountains and in the northern parts of Sweden and Finland. This conflict arises because the relationship between winter temperature and precipitation differs between the long-term climate change and the interannual variability. As discussed below based on Fig. 11, the projected long-term increase in winter precipitation is in most of northern Europe smaller than the projected warming together with the interannual regression relationship between temperature and precipitation anomalies in ERA5-Land would suggest.

The EURO-CORDEX RCMs simulate, on the average, a NDJFM mean warming of ca. 3-5 °C from 1981/82-2019/20 to 2059/60-2097/98, with a general increase from southwest to northeast (Fig. 11a). The change in precipitation varies from slight local decreases in western and northern Norway to increases of up to 25 %, with a relatively sharp northwest-to-southeast contrast across the Scandinavian mountains (Fig. 11b). This contrast is qualitatively similar to that found by Räisänen and Eklund (2012), but its connection to the atmospheric circulation in the EURO-CORDEX RCMs would require further investigation. The multi-RCM mean changes in the NDJFM mean sea level pressure in northern Europe are small (from 0 to +1 hPa), implying only very modest changes in the average lower tropospheric winds (not shown).

The ratio between the precipitation and temperature changes is mostly 2-6 % $(°C)^{-1}$, but lower in western and northern Norway (Fig. 11c). On the interannual time scale, however, a 1 °C positive temperature anomaly is statistically accompanied by a 12-15 % precipitation anomaly in western Norway (Fig. 11d), where westerly flow anomalies result both in advection of warm Atlantic air and forced ascent uphill the Scandinavian mountains. The interannual regression coefficient (Fig. 11d) also exceeds the long-term precipitation-to-temperature change ratio (Fig. 11c) in Finland and northern Sweden. For example, in the grid box closest to Sodankylä, the long-term change ratio (3.4 % $(°C)^{-1}$) is only half of the interannual slope (6.1 % $(°C)^{-1}$) in ERA5-Land. The interannual regression coefficients in the EURO-CORDEX RCMs agree generally well with ERA5-Land (not shown).

As indicated by the last sentence, I plan to omit the original Fig. 10e, because this panel was not central to the main message.

Line 426: “colder areas” Is it possible based on this study to quantify/define what“colder areas” are? Wouldn’t that be of interest for many, when a local climate enters a new regime where variations in winter precipitation is no longer the main driver?

This comment is addressed in the new Figure 8 (and related text) that was included earlier in this response letter.