Response to Referee 1

We thank R1 for this detailed review, which enabled us to significantly improve our article. Enclosed please find a detailed explanation of the revisions we made based on R1's comments. For convenience, comments are in bold and our responses are in italic. Revisions made in the manuscript are presented in italic with grey background.

The work by Verfaillie et al. presents a comprehensive analysis of past and future snow conditions at a mid-elevation mountain range in the French Alps. The regional SAFRAN reanalysis and bias-adjusted RCM experiments covering three greenhouse gas emission scenarios are used to drive the Crocus snowpack model, and model simulations are compared against observations at a single measurement site. The snowpack model is employed in a multiphysics ensemble approach which allows for an assessment of the contribution of snowpack modelling uncertainty to the overall projection uncertainty. Results for a range of snow indicators are presented. Concerning the overall future degradation of the snowpack they largely confirm previous works, but also provide a number of new and useful insights that are at least valid for this specific case.

Overall, I consider the paper as a relevant and interesting piece of work. The methods and data used are comprehensively described and are well introduced (except for the downscaling and bias-adjustment method ADAMONT, which is however explained in detail in a previous paper). The methodological approach is sound and valid. The introduction and the discussion properly refer to existing works in this field, and the conclusions are well based on the results obtained. There are no language issues, and the topic clearly fits into the journal's scope. As such, I could generally recommend a publication of the work. However, a few minor and one major issues remain, and I'd suggest to ask the authors for a revision of their work in these respects before final publication. Minor issues are listed at the end of this review.

The remaining problem with the existing manuscript is its rather technical touch and the wealth of information that is presented in terms of data sets, emission scenarios, scenario periods, methods and especially snow indicators. The comprehensiveness of the work is impressive, but the reader very easily gets lost in this large amount of information that is presented in the text, in the tables and in the figures. These information might be very useful for local stakeholders operating in this very region and being affected by snow conditions, but for a truly useful contribution to the scientific community the results and their presentation need to be much better streamlined in my opinion. The generally most interesting part of the work is probably the entire methodological approach and the possibilities that arise from it. The very detailed results for a representative elevation of 1500 m in the Chartreuse mountain range are more of a case study, and the details of their presentation should receive less emphasis. One option might be to remove parts of the analysis entirely from the manuscript and place it in an additional, accompanying publication (e.g. the multiphysics ensemble analysis which is only briefly described in the results and which has much more potential to be analyzed in more detail). Another option is to move some of the material from the main manuscript to the supplement. This could for instance concern several of the snow indicators (like onset and meltout date), that are in any case only briefly discussed.

We thank R1 for the overall positive appreciation of our work and for the suggestions for improvements. Following them, we have decided to move Tables 3 and 5 about STEDs to the supplement as those are only briefly discussed in the text. This will reduce the total number of tables in the main article. We acknowledge the wealth of information provided in this manuscript, and consider worthwile to present both the methods and scientific results from an examplary

geographical configuration together. We also believe that the geographical location considered (Chartreuse, 1500 m altitude) has broader relevance and that conclusions reached for this case are worth being presented in detail. We agree that several aspects of the present work deserve more in-depth analysis, and this could be addressed by future publications, although some key features can already be analyzed and assessed from the present manuscript

While we acknowledge that the content of the manuscript can indeed be considered rather dense, we value this to be a positive quality judgement rather than an issue for a scientific publication, which targets a specialized audience.

We also hope that the dense content of our manuscript will be easier to follow after several clarifications in the text. For example, the geographical setting of the case study is now better introduced and separated from the description of the observational datasets (see below). The description of the statistical post-processing was also clarified (see further).

In combination with such a streamlining, I'd suggest to put a little more emphasis on the actual processes that are responsible for the identified future changes in the snow indicators. Little is so far said about that. The Crocus model output surely provides an opportunity to do so (e.g., separate analysis of snow accumulation and snow melt amounts).

This was already approached in our paper (through the analysis of temperature and precipitation, including the phase of precipitation, as the main drivers of snowpack reduction). While interesting, a further analysis seems out of the scope of our paper, and would make it longer.

In this respect, the relation of the snow indicator changes to the GLOBAL temperature change is not very helpful and the authors should think about putting the LOCAL temperature change into focus (though I completely understand the choice of the global scale given political climate targets).

Local temperature changes are addressed in detail in the manuscript, in text, tabular and graphical formats. We acknowledge that the reviewer understands our choice to relate local changes in snow indicators to global temperature variations. The reasons and the limitations of this choice are detailed in section 4.4. We understand the reviewer encourages us to relate local changes in snow indicators to local changes in temperature. While doable and certainly leading to significant relationships (at least on 30 years average values) because the physical link is obviously more direct, we preferred not to include this in the revised manuscript in order to not lengthen it and not induce confusions between local temperature and global temperature relationships. This could be addressed in a future study, either by us or other research groups, which may be interested in exploring further results, which could be obtained on the basis of our newly derived dataset..

Such a shift of the focus away from details of the case study and towards a more methodological and process-oriented analysis would be very worthwhile in my opinion. Apart from this and as said before, I consider the manuscript as being of high quality and of general relevance for the readership of the journal.

We understand the reviewer's point of view, and consider this manuscript to be viewed both as a methodological and application oriented manuscript. We will be pleased to introduce future publications targeting expanded application domains (entire French Alps, Pyrenees, etc.) as well as more in-depth analysis of the drivers for snowpack changes.

Minor issues ===== Spatial scale of the Crocus application: What remains somehow unclear is the spatial setup of the Crocus model. I assume the authors use a single-site setup, driven by the outcomes of the SAFRAN reanalysis and of the ADAMONT downscaling method for a representative 1500 m elevation range in the Chartreuse massif. Is that the case? If so, this needs to be clearly said and described in some more detail. It would imply

that the results shown are only valid for that specific elevation range in this massif. What about other elevations then? Is it possible to come up with some speculation here as well? Snow projections will surely strongly depend on the elevation considered, and some placement of the results into a broader spatial context would be helpful.

Yes, we use a single-site setup (Chartreuse massif at 1500 m as.s.l.). This is now better explained through a new section « 2.1 Geographical setup ». We preferred not speculating on results which could be obtained at other geographical locations and altitudes, because this would unnessarily lengthen an already long manuscript, and will be addressed without speculation in follow-up publications. However, we added the following in the Conclusions : « (…) our results do not directly allow extrapolation of the conclusions in other mountain regions in France or other elevations, although it is expected that the response of neighbouring mountain ranges may be comparable at the same altitude level. ». (p. 30 L. 25-27)

Page 1 Line 2: "investigates" instead of "introduces" is probably the better choice.

This is now corrected. (p. 1 L. 2)

Page 1 Line 9: "reduction in mean interannual snow conditions" is rather unclear.

We have rephrased : « reduction in average snow conditions ». (p. 1 L. 9)

Page 2 Line 30: "they" instead of "there".

This was corrected. (p. 2 L. 29)

Page 3 Line 22: The term "currently" is probably wrong. At this point, more GCM-RCM chains are available from EURO-CORDEX. The authors just either specify their date of access of the data base or justify their selection of all available model chains.

We used the models available when we last accessed the database to retrieve the data before launching the whole processing chain, and for which the geopotential data for the corresponding CMIP5 GCM (necessary in the ADAMONT method for the calculation of weather regimes) were available.
This is now specified : « The 13 GCM/RCM EURO-CORDEX pairs available in April 2017 (and for which the geopotential data for the corresponding CMIP5 GCMs were available) were used. These are expected (...) ». (p. 3 L. 21-22)
And also in Section 2.3 : « This study uses the EURO-CORDEX dataset (Jacob et al., 2014; Kotlarski et al., 2014) available in April 2017, consisting of (…). Only the GCM/RCM pairs for which the geopotential data for the CMIP5 GCMs were available were used. » (p. 5 L. 2-5)

Page 4 Lines 8-15: Please check: Is SAFRAN really ONLY available over mountain ranges? To my knowledge, entire France is covered.

SAFRAN refers here to the original mountain region implementation (Durand et al., 1993). SAFRAN was expanded to wider geographical areas in France (Vidal et al., 2010) and Spain (Quintana-Segui et al., 2017). This is now indicated. (p. 4 L. 21-22)

Page 8 Lines 5-21: This method description is rather confusing and very hard to follow. Please streamline.

We are sorry that the statistical post-processing appeared confusing, indeed this is a key aspect ouf our work and we attempted to make it clearer in the revised manuscript, which now reads (entire paragraph copied here, p. 7 L. 10-33) :

« The entire model chain provides estimates of a series of annual indicators spanning continuously the historical period from 1950 to 2005, typically, to the end of the 21 st century. A total of 13

GCM/RCM pairs were considered in the case of RCP4.5 and RCP8.5, out of which 4 are also available for RCP2.6. We generally used a 15-year window to assess the statistical distribution of the indicators considered. For a given GCM/RCM pair and a given RCP, statistics corresponding to a given year can be computed using indicator values for the 15 years surrounding it (7 before, the central year, and 7 after). In what follows, we assume that all GCM/RCM pairs bear equal probability (Knutti et al., 2010). We post-processed the distribution of annual indicator values in two ways.

1. Quantiles of annual values: In this case, for a given RCP, all annual values of the indicators spanning the 15 year time window for all the corresponding GCM/RCM pairs were pooled together (195 in the case of RCP4.5 and RCP8.5, 60 in the case of RCP2.6). The quantiles of the distribution of the annual values were determined using a kernel smoothing approach. We computed the 5%, 17%, 50%, 83% and 95% values (Q5, Q17, Q50, Q83, Q95), consistent with IPCC (2013). This approach provides statistical estimates for annual values of the indicator, although it mixes together the effects of interannual variability and inter-model variability.

2. Moments of multi-year averages: A running average of annual indicator values was computed using the 15 year sample window, for a given RCP and for each GCM/RCM pair. For a given RCP, mean (μ) and standard deviation (σ) values were computed for the ensemble of multi-annual averages of all GCM/RCM pairs. This approach provides information on the statistical distribution of each indicator for a given RCP on a multi-annual average perspective. In practice, we compute σ' = 0.95 σ, corresponding to the 17% and 83% quantiles in the case of a normal distribution, so that this approach becomes more comparable to the annual quantiles approach described earlier. In the case of the multiphysics Crocus model implementation, we mostly used the multi-year averages approach, and applied it to all Crocus members.

The spread of the distributions of these two approaches can be assessed in rather similar ways. In the multi-year average approach, the coefficient of variation CV can be determined as CV= 2 × σ 0 /μ. In the annual quantiles approach, the spread can be assessed by dividing Q83-Q17 by Q50 to form a formal equivalent to the coefficient of variation, defined using quantile values instead of mean and standard deviation (referred to as quantile-based coefficient of variation -QCV- hereafter). ».

Figure 1: The STEDx should represent some duration of exceedance and hence need to be represented by some horizontal range in this graph. The representation by single vertical arrows is probably wrong, please check.

We agree with this remark and have attempted to improve the graphical representation of this series of indicators.

Page 9, line 1: Temperature changes are surely not computed in a relative manner, please check.

OK. We changed the sentence to : « Changes were computed (...) ». (p. 9 L. 11)

Page 19 Lines 6-10: I assume this is simply an effect of random internal variability at decadal scale, could that be (simulations out-of-phase with reality)? Please clarify.

We agree that this is an effect of random internal variability. This is what we wrote in the original manuscript (« low frequency variations at the decadal time scale, superimposing on a long-term trend of general snow reduction »).

Page 23 Line 31: Isn't it rather random variations (instead of systematic variations)?

We removed the word « systematic ». (p. 25 L. 5)

Page 25 Lines 14-16: Is this really the case? Why should a matching of quantile distributions reduce interannual variations? please check and better explain.

The method does not reduce interannual variations, but it is clear that it will tend to reduce the spread between different GCM/RCM model results over the calibration time period. We have rephrased to : « which inevitably reduces the spread between different GCM/RCM pairs ». (p. 25 L. 22)