Second review of "Snow sensitivity to climate change during compound cold-hot and wet-dry seasons in the Pyrenees"

Dear authors, dear editor,

The revised version of the manuscript addresses some of the points raised in the first round of review. The method is better detailed, a section on limitations and uncertainty has been added, and the language has been improved. However, I still have some concerns. One is about the vocabulary used, which can easily be corrected, and a more fundamental one about the method, or rather how the results are discussed, which was already raised in the first round of comments and remained unanswered. While this may require more work than was done for the first iteration, I really encourage you to address this issue to increase the robustness of your results.

Best regards,
Adrien Michel

Important note: The updated manuscript does not agree with the track change version! E.g. P2L36 " … and **increasing** surface and air temperature …" in the manuscript, and " … and **reducing** surface and air temperature**s** …." in the track change version. The last sentence of the conclusion also differs (I did not check further). My comments are based on the updated manuscript. The author should carefully check on which version they are working on for further edits.

Vocabulary

There are been many improvements in the usage of the term "sensitivity" compared to the first version. However, in many locations, the word "sensitivity" is still used alone, while "sensitivity of snow to climate change should be used". E.g. section 4.2 is called "snow sensitivity", while it should be "snow sensitivity to climate change", or more precisely "Snow sensitivity to change in air temperature and precipitation".

In the text, you mainly use "extreme compounds seasons", while in the title you use "compound cold-hot and wet-dry seasons" only. While the term "extreme" could be disputable here (based on percentile 60), this is not unique in the literature (despite usually higher percentiles are used), so it is fine. However, for consistency you should maybe also use "extreme" in the title.

The term "ablation" is often used alone, while I think you are most of the time referring to "ablation rate", this should be corrected. Also, it is not clear if you talk about absolute ablation rate (cm/day) or relative ablation rate (in %/day or %/°C). Both of them are relevant, but it should be clarified to which one you refer to.

Analysis description

You added some details on P6L170-178. However, we still do not know exactly which model setup you used. If I want now to reproduce your work, I need the exact name of the models' parameters. I Imagine these parameters are chosen at compilation time form the compile.sh file (https://github.com/RichardEssery/FSM2/blob/master/compil.sh):

```
#define ALBEDO 2   /* snow albedo                   : 1, 2     */
#define CANMOD 1   /* forest canopy layers          : 1, 2     */
#define CANRAD 1   /* canopy radiative properties: 1, 2     */
#define CONDCT 1   /* snow thermal conductivity    : 0, 1.    */
#define DENSTY 1   /* snow density                  : 0, 1, 2  */
#define EXCHNG 1   /* turbulent exchange            : 0, 1     */
#define HYDROL 1   /* snow hydraulics               : 0, 1, 2  */
#define SNFRAC 1   /* snow cover fraction           : 1, 2     */
/* Driving data options                   : Possible values.   */
#define DRIV1D 1   /* 1D driving data format        : 1, 2     */
#define SETPAR 1   /* parameter inputs              : 0, 1     */
#define SWPART 0   /* SW radiation partition        : 0, 1     */
#define ZOFFST 0   /* measurement height offse      : 0, 1     */
/* Output options                         : Possible values    */
#define PROFNC 0   /* netCDF output                 : 0, 1     */
```

A table in supplementary with the exact names (and maybe the exact version of the model, i.e. git commit number) is necessary for reproducibility.

There is no "data and code availability" indicated in the paper. I highly encourage you to publicly share your data.

Impact study, determining factors, uncertainty

On P11 L284-285 you say: "The results show a non-linear response between seasonal HS loss and temperature increase." Which is clear from figure 4 and an interesting result. However, later in the analysis, you mainly use linear indicators (in %/°C): e.g. P11-12 L292-319, Tables 2-3. I think these numbers do not add anything to the analysis and they contradict the non-linearity you found and emphasis in your abstract (saying that the greater relative change is for +1°C). I would recommend to remove the above-mentioned lines and tables. This will make this Section easier to read, without any loss of information. The same information is obtained by commenting the boxplots on Figs 5-6-7. As I already mentioned in the first review round, there are many numbers listed in the text, which are all visible from figures, and not really useful later on in the analysis.

In the first revision round, I raised this important concern:
"'I have one concern about the method itself. As far as I understand, seasons "classes" (WW,CW,etc.) are determined for each subregion and elevation range separately (Figure S1). And thus, figures like 4 are obtained by averaging all the regions together for each elevation band and season class. **My problem is that from Figure S1 we see that some classes of**

season are manly dominated by some regions (e.g. cold wet is dominated by south-west regions). So, when comparing the different season class, we do not really know if the difference is due to the meteorological input, or due to some other aspects differing between regions. In addition, the season class is (maybe?) determined for each region separately (see my comment above), so a CW in one region might not be CW in another region. **As a consequence, because of the approach chosen, I do not think the differences observed between compound seasons is only due to the specific weather of the seasons. This is probably the dominant factor, but the spatial difference would add some uncertainty there. This should at least be discussed**."

Many points have been answered and clarified, but not what is in bold font in the paragraph above (you just answered "We are not comparing season types between massifs"). I'll reformulate here this concern.

In Figures 4 to 7, and in most of the analysis, you split the data per elevation range and per season type. In Figure S1, we see that some seasons types occur mor often in some regions (e.g. some regions in the south have no warm-wet, some regions have a total of ~80 extremes seasons, while some have only ~40 in total). As a consequence, when looking at one class of season and one elevation band, the different regions are not represented equally, on some seasons type signal will thus be influenced by the dominant regions in the sub-ensemble (e.g. col-wet mid altitude are dominated by the western regions). In Figures 9-10 you show that the response to climate change is different between region. So, when in the end you assess the change for a season type and an elevation band (e.g. col-wet mid altitude), we do not know if the response is more dependent on the season type, or on the region which is dominating this subset. Looking at the spread in the boxplots of Figure 6, I'm not sure that for all case we have a proper statistical difference between low and mid elevation band (this can be statistically tested). This can be explained by the fact that going from low to mid altitude the representation of the regions in the sub-ensemble considered in not the same. In other words, we cannot with certainty attribute the difference to the season type as you do in the analysis. Note that you can do some statistical tests to see if the different response to climate change between region, elevation band, and season-type are significant, excluding the two other parameters (note: using only one variable you would not have this problem, because by definition all region will have the same number of extreme, 40% of the seasons with the percentiles you use).

This imbalance between regions, induced by the joint quantile approach used, should be discussed in the text (note: it indeed totally makes sense to compute extreme per regions). This is not an insurmountable problem, but this is a clear drawback of the method used (as every method has). A clear example is the following sentence in the conclusion: "In particular, snowpack losses were greatest during WW seasons at low and mid-elevations and were greatest during WD seasons at high elevations". At mid-elevation, the eastern region has more WW event that the western one (Figure S1), at the same time, HS is more sensitive to climate change in the eastern regions (Figure 9). Now the question is: Is eastern more sensitive because of the local conditions (e.g. closer to isothermal conditions in the baseline simulation), and thus since it has more WW season, the WW signal will appear more sensitive simply because it contains more seasons from this region? Or is it the opposite: WW season are for some reason more sensitive to climate change, and eastern region having more WW season compared to the other regions, it appears to be more sensitive to climate change? Are the local conditions or the season types dominant here? The fact is that with this analysis we cannot answer this question. We can see some correlation between season and sensitivity to climate change, yes, but we cannot attribute the observed different sensitivity to the season type, this is a major difference.

The extreme compound season is really emphasised in your abstract/title/conclusion, but:

1. The analysis suffers from the problem discussed above. We cannot do a proper attribution.
2. Is not that much discussed in the text in the end. Indeed, only one paragraph (Section 5.3) discusses it in the whole discussion. Finally, most of the discussion is based on more general results.

I think here lies my main problem with the current status of the manuscript. Either the focus is kept on the compound seasons (then maybe the general discussion on well know impacts of climate change on snow (Section 5.1 and 5.2) should be highly shortened), and the strength of the analysis on the difference in the signal between seasons should be improved (by using proper statistical test showing that season type is significant despite the imbalance in regions) and the problems mentioned above arising from the season type construction need to be discussed (which I encourage you to do); or you decide to be more generalist (as you are now in some parts of the discussion), and then you remove most of the emphasis on the compound season.

Minor comments

P1 L35: Should be “increasing” the albedo.

P2 L36: Should be “decreasing surface and air temperature”. And it is not absolutely true that snow decreases surface temperature. During winter, the snow/soil interface will mostly remain at 0°C if the soil is snow covered, while if the soil is snow free but the air temperature is cold, the soil surface temperature will further decrease (snow is a really good insulator for the soil). In spring, it is true that snow cover will keep the soil colder. However, I would keep only “decreasing air temperature”.

P2 L59: “on snowpack duration” → “on snow cover duration”

P3 L85-86: “and the different mountain exposure to the main air masses”. Which “main air masses”?

P4 L95: What is “mid-late”?

P4 L105: “Sensitivity” to what? (Same at lines 107, 108, 110, Sections 3.4 and 4.2 names, etc). Should always be “sensitivity to climate change”, this is indeed a bit heavier in the text, but correct (see comment above).

P4 L115: What does “these” refer two? Long sentence with many commas, hard to follow, consider splitting in two sentences.

P9 L259: Do you mean “ablation rate”?

P9 L262: Why isn't snow duration also an accumulation indicator?

P11 L287-288: "High elevation areas had lower season-to-season snow variability than low elevations for all season types (Figure 4)". I don't see any information about season-to-season variability in Figure 4.

P11 L 289: Avoid using the word "significantly" if not in the context of a proper statistical analysis (and thus a statistical "significance").

P11 L289-291: "All the snowpack-perturbed scenarios indicated that snowpack decreased at low and mid elevations under warming climate scenario". This is also the case at high elevation (Figure 4).

P13 L331: "the peak HS date per °C was earlier by 9 days". This does not mean anything to me, should be: "the peak HS was anticipated by 9 days per °C" (but see my comment on linear indicators).

P13 L337: "and because": Remove "and"

P13-14 Figures 5-6: Why not having only one figure with three rows of boxplot?

P14 L350: "At low elevations, the snow ablation in all four extreme seasons was 12%/ºC". Something is missing here. Should be "snow ablation rate increase". Same for the rest of the paragraph, should be ablation rate change/increase/decrease.

P16 Figure 8: I think the y-axis units should be cm/day. Explain the numbers in the caption, e.g. "the numbers in the plot show the difference in ablation rate compared to the previous degree". Or maybe I don't understand the figure, and this just are absolute values of ablation rate. So then why stacking them on top of each other and why having a y-axis? Shouldn't it be like figure 7? Also, I can't reconciliate what I shown in this figure with the numbers in Table 4 for the ablation columns. In the text and Table 4, you have ablation in %/°C, and here in cm/day.

P18L398: "The sensitivity of snow to different spatial patterns of climate change that we identified here […]". You do not study different patterns of climate change; you apply the same delta everywhere. Please correct.

P21L504-505: "Our maximum snow ablation and peak HS date occurred during dry seasons …]". Are you talking about change or about absolute value in the baseline simulation?

P21L508-510: "The temperature in the Pyrenees is still cold enough to allow snowfall at high elevations during WW seasons, and for this reason we found maximal sensitivities during WD seasons." I don't understand, temperature is almost the same in WW and in WD, so why is sensitivity to climate change greater in WD?

P23L562-564: "however, a more complex model does not necessarily provide better performance in terms of snowpack and runoff estimation (Magnusson et al., 2015)". But it also can, especially for climate change study (see Carletti et al, 2022, doi.org/10.5194/hess-26-3447-2022)

P564-566: “Biases in the SAFRAN system and biases related to the FSM2 were minimal because we quantified relative changes between a modeled snow scenario (climate baseline) and several perturbed scenarios”. This assumes constant biases. Snow cover involves different variables, non-linear processes, and will accumulate errors along the season, we cannot be that certain.

P23L568-569: “[…] but assumes that the snow patterns of the reference climate period will be constant over time.” Don’t you mean meteorological patterns? (E.g. you don’t capture change in precipitation regime, you just scale the intensity of each events). Please clarify.