We thank J. Parajka for his encouraging and constructive comments. We address the reviewer‘s issues (in italics) below. Changes in the manuscript are given in blue, answers to the reviewer are given in black.

1) *The main question I had when reading the manuscript is to what extent it is possible to make robust and strong interpretations from using only 2 images (in different regions and years). In my opinion, many statements are too strong and need to be validated (supported) by using much larger number of images, describing different snow conditions (snow poor/rich winters, different periods of snow season). For many practical applications, the temporal variability/uncertainty of snow cover sub-grid parameterization might be more important than one fixed relation found in one winter maximum. I would thus suggest to very carefully and critically consider the limitations of using just selected examples of snow distributions and to revise some interpretations made. Some statements seem to be too general and are likely not fully supported by presented results.*

We clearly agree that other spatial snow depth data sets are required for an even stronger validation of our derived parameterizations. In the revised manuscript we hopefully point out the current limitations with more clarity than in the original submission. We went through our statements and revised many of them to point out the limitations even more (e.g. we completely revised the Discussion and Conclusion section).
However, we want to point out that the inclusion of the Spanish data set was our first step towards a development of a parameterization independent of one geographic region. It is based on a different snow climate and different winter compared to the two data sets in Switzerland. So far we had to work with peak of winter data. Based on a previous result of observed annual persistent snow depth distributions in the same areas at peak of winter (e.g. Schirmer et al., 2011) our goal was here to test if the snow depth distributions can be described by topographic parameters alone and if different snow climates (or equally different snow conditions) can be brought in by applying the current mean snow depth. Our correlation analyses of the standard deviation of snow depth with terrain parameters on different scales show similar trends and magnitudes for all three areas (Fig. 5 and 6) and therefore confirmed our hypothesis. Given the limitation to two different snow climates our climate/seasonal indicator has to be re-evaluated once new highly-resolved snow depth data sets become available. However, new data will likely only change the constant parameters in Eq. 2. We believe that besides these obvious limitations our approach provides a good description of the dominant processes on larger scales than only a few hundred meters.

2) *Some of the expressions (terminology) might be misleading. In the motivation, there is expressed a clear need for proper sub-grid parameterization of snow cover for climate and regional modeling, (and I agree with it), however the typical grid sizes of regional models (5-50km, even more for large-scale climate models) goes beyond the upper limit of grid sizes tested in the manuscript (3km). So I wonder if are the results directly comparable and applicable for studies using coarser grids? Please consider this point when introducing the objectives and discussing the findings with existing applications of regional (and climate large-scale) models. Is the relative role of sub-grid variability so important also in e.g. 100m grid snow modeling?*

We believe that our parameterization is applicable for grid cell sizes comparable to those of large-scale meteorological and regional climate models. Due to a lack of highly-resolved snow depth data in regions with larger horizontal extensions it was however not possible to test larger grid cell sizes than the 3 km. In our regions, we would have obtained too many similar domains from our random sampling procedure. However, we did perform a scale analysis. Overall, we found less scatter the larger the grid cell size (or domain size L) and thus higher correlations (e.g. Fig. 6). Our parameterization for the snow depth distribution also shows higher accuracies the larger the domain size. The reason for this is, that with increasing domain size the dominant subgrid topographic features can be more accurately captured ($L \gg \xi$) also leading to a smaller correction term for finite grid cell sizes (third term in Eq. 2). On the other hand, the decreasing scatter with increasing domain size reveals the diversity of processes shaping the snow depth distribution at small scales. At larger scales the large-scale topographic influences on precipitation and on the shortwave radiation balance, which were described by the second and third term in Eq. 2, dominate. Describing subgrid variability for grid cell sizes of 100 m is rarely feasible (see scatter in Fig. 8a) and was not the scope of our study. The smallest grid cell sizes were included to demonstrate the scaling of snow depth distributions starting from the measured values in 2 m horizontal resolution. Thanks for pointing out that unclarity. In the revised manuscript we now make our intentions more clear (revised Introduction and Discussion and Conclusion).

3) *Having said that, I would suggest to strenghten the story of the paper. Some more detailed outline in the introduction would be helpful to better understand, why is the scaling analyzed first, snow cover sub-grid parameterization later. Please consider also, to more clearly demonstrate the advantage and reasoning for using airborne data for deriving sub-grid snow cover parameterization, in comparison to other methods. The benefits of using airborne data are not clearly formulated and discussed.*

Thanks, we agree and therefore thoroughly revised the introduction incorporating your points.

Specific comments:

1) *The selection of three different regions is not clear. Why Spain in a different winter? The size around 30km2 seems to be rather small for making robust interpretations for coarser grid sizes.*

Two highly-resolved snow depth data sets covering large areas were gathered with an opto-electronic line scanner. To develop a parameterization independent from one specific climate we needed more data sets. Unfortunately, so far we neither have access to data from repeated flights nor other areas covered. The data set in Spain gathered by airborne laser scanning data covers the largest, coherent area (aside from our Swiss areas) we had access to and shows a dryer snow climate than in eastern Switzerland. Hopefully, we will have more snow depth data sets in the future covering similar extensions and showing similar horizontal resolutions. We already discussed the overall smaller size of the Wannengrat and the Val de Nuria areas in the Method section 3.1 as well as in the Discussion and Conclusion section (second paragraph).

2) *The Summary and Discussion section reads really as a summary only. Linking the findings with existing literature (in a separate section) will help to more clearly indicate the benefits and challenges of using airborne data for sub-grid snow cover parameterization.*

Thanks for pointing that out. We have thoroughly revised the Discussion and Conclusion section and hope that it reads better now. However, we like to keep the one section for the Discussion and Conclusion.

3) *Figure 1. It would be interesting to present also snow depth distributions used for the analyses.*

In Figure 2, we already showed the snow depth distributions used for the analyses.

4) *Figure 3,6,8: Please consider to use a discrete color legend instead of continuous.*

We now use a discrete legend in Figure 3, 6, 7 and 8.

5) *Figure 9: It would be interesting to see also some validation of derived snow cover depletion curves.*

We agree, but unfortunately we do not have available spatially measured, time-dependent snow depth data sets.

We thank D. Bavera for his encouraging and constructive comments. We address the reviewer's issues (in italics) below. Changes in the manuscript are given in blue, answers to the reviewer are given in black.

1) *I would better explain and describe the selection procedure of the test sites where the method has been applied and possibly I would add a few more case studies in order to give more generality to the paper outcomes*

Unfortunately, we could not select from a variety of data sets since large, coherent snow depth data sets in a high resolution are rare and not always publicly accessible. We described our available sites and the measurement methods in section 2, in the original manuscript.
It is therefore also not possible to add more case studies since we do not have access to other highly-resolved snow depth data sets at peak of winter or during the depletion period covering large regions comparable to the size of the three we used (larger or equal to 30km^2). Having coherent regions of highly-resolved snow depth data that cover elevation differences larger than 1500m and being large enough to randomly sample in differently sized domains (see Fig. 1) is quite unique.

2) *Some of the conclusion appear too strong and too general considering previous observation and moreover that it has been applied only for peak of winter for only a small number of snow conditions*

Thanks for pointing that out. We revised the Discussion and Conclusion section (see also our response to the first reviewer). We now also clearly state that our three data sets were all gathered at peak of winter in two different geographic areas. However, we do not fully understand what is meant by 'a small number of snow conditions'. Our parameterization was tested in two alpine regions with different topographic characteristics. However, since we lack large-scale measurements from other environments, such as prairies, tundra and forested regions, our parameterization is restricted to alpine regions. Once more highly-resolved, large coherent snow depth data sets become available for different regions the application of mean snow depth as a climate indicator can be verified. In the last paragraph of the Discussion and Conclusion section we now discuss this issue.

3) *I would better highlight the limit of the paper outcome related to its data (site selection, time, variability of snow conditions)*

In the revised Discussion and Conclusion, we hopefully now better point out the limitations in terms of the data sets which we had available. See also our answer to the previous issue.

4) *Provide a more clear interpretation from the physical processes point of view of the paper outcomes in order to better understand the meaning and the relevance of the results deeply and precisely described in section 4*

We revised the Discussion and Conclusion section regarding the physical processes discussion.

Specific comments:

1) *Please better clarify how you individuated peak of winter timing*

We here define the peak of winter as the point in time with the highest snow depth during each individual winter. Thus, a measurement had to be conducted as closely as possible at this point/period in time. However, obviously it is very difficult to choose a measurement date which actual falls precisely at this point in time since the actual course of the winter is not known in advance. Please also keep in mind that high-resolution snow depth observations from aerial photography and from ALS require cloudless sky conditions. For more details about the data acquisitions see Section 2 in our manuscript, where also the articles describing the measurements in detail are cited.

2) *At line 19 p. 9792 describe under which hypothesis you assume spatially homogeneous melt*

The assumption of homogeneous melt rates might be questionable, even though it was previously found that homogeneous melt rates reasonably depict a depleting snow cover in mountainous terrain; see the study of Egli and Jonas (2009) which we cited in line 13 on page 9806. We used this assumption to derive snow cover depletion curves in order to verify a previously published SCA parameterization which was also derived assuming homogeneous melt rates (cf. Essery and Pomeroy (2004)).

3) *At line 24 p. 9797 subtraction should be the opposite. HS = winter-summer*

Thanks. Of course, HS was derived by subtracting the summer from the winter DSM. We corrected the typo.

4) *If possible please better motivate and describe the differences between the three study sites shown in table 1*

We feel we described the correlations with terrain parameters for all regions and each catchment separately quite sufficient in the original results section (on p.9803 lines 15-26) as well as in the original Discussion and Conclusion section (on p. 9809 lines 15 to 25).

5) *I appreciated the readability of figures and plots*

Thanks.