We thank the reviewer for their encouraging and constructive comments. Their comments (in italics) are addressed below. Answers to the reviewer are given in blue.

Major comments:

The language needs to be improved to be more concise. Just as one example: P1L17: Would snow in another season not be intercepted? Both in this sentence and the next one I assume the authors mean that in a coniferous forest 60% may be intercepted. As it reads now, 60% of some total are intercepted in coniferous forests and 24% are intercepted in deciduous forests in the Andes, i.e. 84% are retained in total. I agree that this is a minor detail and one can guess what the authors mean, but in a scientific paper these things should be formulated as clearly as possible.
The manuscript was carefully read by all co-authors and additionally by several native English speakers from the USGS. The language was therefore carefully checked before the submission. Nevertheless, to avoid any unclarities in the manuscript, we will carefully go through the manuscript once more, and make the wording more concise where appropriate.

Central parts of the methods are described first in the result section.
Unfortunately we do not understand which parts you mean. All methods are described in the methods section. The result section is structured as follows:

4.1. Grid cell mean snow interception
4.1.1 Parameterization
4.1.2 Validation

4.2 Grid cell standard deviation of snow interception
4.2.1 Parameterization
4.2.2 Validation

The resulting parameterizations are not part of the method, they were developed in this study using the data given in the data section, the forest structure metrics and the method both described in the methods section (3.1 and 3.2). As such the parameterizations (i.e. 4.1.1. and 4.2.1) are part of the results. The validation sections describe how modeled interception compares to observed interception (i.e. 4.1.2. and 4.2.2).
While we do not see that parts of the results should be moved to the methods we agree that the headings could be more concise and we will consider changing them.

The field observations need to be described in more detail. I honestly do not understand what has been measured how. It also sounds as if some data were selected from a larger set, the reasons for this are not entirely clear.
Unfortunately, we do not understand which data you think were selected from a larger set. However, we will carefully go through the data section to clarify the description of the measurement methods.

The two central equations suddenly pop up in the result section. How were these two types of equations derived? Is there any physical reasoning for certain functional relationships like the exp or power function? How exact can the coefficients be determined? Uncertainty? Sensibility?

In the method section
3.2 Subgrid parameterization for forest canopy interception
we describe that the empirical parameterizations are derived using the Swiss data set (Line 211 ff).
All parameterizations were empirically developed using the Swiss development data set. The existence of varying previously observed functional relationships (base functions) were mentioned in the introduction (Line 43-46) and were considered here. In the results section (Line 243-246 and 251-259) we explain the reasoning for our functional relationship. In the discussion (Line 389-405) we largely discuss our choice compared to previously published base functions.
To specify the robustness of our coefficients we will include the confidence interval of the coefficients.

Furthermore, I do not understand what the stdev of the DSM is. Variation of ground surface? But this would not have anything to do with the trees. Variation of vegetation heights? But then DSM is the wrong term.
Forest structure is parameterized here from the Digital Surface Model (DSM) which is the top of the surface, i.e. surface elevation + vegetation height (DTM+CHM). DSM is a standard abbreviation for the surface height (e.g. https://www.earthdatascience.org/courses/earth-analytics-python/lidar-raster-data/lidar-chm-dem-dsm/). The abbreviation DSM was introduced in line 92.

The standard deviation of DSM σ_z describes the variation of vegetation altitude by integrating both the variability of the canopy height and of surface elevation. Using σ_z seems more realistic because gaps and spaces between trees are influenced by local topography. The standard deviation of DSM σ_z was introduced in the methods section as our second forest structure metric (3 Methods / 3.1 Forest structure metrics in Line 206-210).

My major concern regarding usability is the choice to express everything as snow height rather than SWE. When used as part of a larger model, I would assume one is most often interested in SWE rather than heights. Also conceptually I am not sure what the height of intercepted snow implies? Height on branches? Probably rather height as the snow would be if being on the ground? But then at which density, that of the other snow on the ground or that of the intercepted snow? Sorry, but I find this very confusing and limiting. Thus, I would prefer to see the interception etc expressed in SWE.
We deliberately chose to parameterize snow depth over SWE because snow depth was spatially measured on the ground and not SWE.
We could have applied an empirical new snow parameterization to derive spatial SWE based on e.g. interpolated air temperatures (as e.g. in Moeser et al., 2015b). This would however have introduced a lot of uncertainty in an interception model since this is determined by the applied empirical density parameterization, measurement errors in air temperatures as well as by the spatial interpolation of temperatures. This was discussed in Line 214-220.
We therefore decided to derive spatial snow depth interception estimates from snow depth observations within and outside of the forest. Using our interception models spatial snow depth in forested regions can be more realistically described. Converting snow depth to/from SWE with a density parameterization and its connected uncertainties is controlled by each snow module (as part of a complex model) and these uncertainties will not be linked with the presented snow interception model.

We will expand our explanation in Line 214-220.

As the two equations are derived from data for ideal situations (no prior snow …) I am not sure how these should be used for the real case, where there is often a history of prior snow on the trees. It seems here one might run into the problem that a simple empirical equation is not really a model after all.

You are right our parameterizations were developed on a data set that had no prior snow on tree branches when the precipitation event started. Nevertheless, both validation data sets did not have this prerequisite but still compared well to modeled interception using the novel empirically derived parameterizations. Especially the interception data set from the US often integrated snow interception over several storms due to longer time periods between data collection. Thus, the trees weren't necessarily snow free for a following snowstorm. Instead these measurements may have been influenced by snow settling, wind redistribution, sublimation, unloading and melt.

All this was discussed in the discussion (Line 359-362).

For a 'model' I would expect some canopy storage accounting, which is an aspect that is missed here.

We have focused on improvements of an interception model rather than multiple related processes. Modeling forest canopy involves several processes, each of which are described with separate models. This includes, unloading, melt and drip (some models), and sublimation. These models use the interception model to dictate how much snow is in the branches at any point in time. If there is still snow in the branches, then it is depleted by the unloading, melt and drip and sublimation models. Thus, canopy storage is dictated by the interplay of each individual model.

In line 361-363 we discuss that we present a model for one forest process. We will add some extra explanation. Furthermore, we will make this clear in the method section: "3.2 Subgrid parameterization for forest canopy interception".