Dear editor Prof. Dr. Markus Weiler,

We want to thank you for thoroughly reading our revised manuscript (RM1). Indeed, the two reviews are "quite different". We, the authors, would say they are contradictory: Reviewer #1 admits some improvements but basically repeats arguments from the discussion phase, although they were addressed (i.e., changed or answered) during the revision. Reviewer #2 grades the paper as "excellent" in significance and quality. Being aware of this complicated situation, we regret your decision to solely follow reviewer #1's argumentation.

You highlighted four points, which we address in the following. Afterwards we comment reviewer #1's statements. The revised manuscript (RM2) was uploaded to Copernicus Publications. We followed most of the suggestions and again made a lot of changes between RM1 and RM2 (shown in green below and marked in the track-change file), but we want to emphasize that none of these changes touches the core of the paper, since Reviewer #1 does not have concerns about the model architecture, the method, the calibration, the validation, the results, the discussion, the plots, the tables.

We are sure the manuscript was improved during this second major revision and want to thank you and the referees for your suggestions. Changes in style, wording and language are partly significant and consulting a native speaking professional was very helpful. As there was no more criticism on the content, the method and the model as such, we are confident that this work now fulfills the high standards of HESS.

Sincerely,

Michael Winkler, Harald Schellander, Stefanie Gruber

Answers to editor points 1) to 4):

1) "The language is still difficult - please try to find a native English speaker to improve the language".
   The manuscript was given to a native-speaking professional. All her suggestions were implemented in RM2. We thank for this work in the acknowledgements.
   Important changes suggested by the native are: (1) the re-writing of the abstract and (2) moving the paragraph on the importance of SWE from the end of section 1.1. to section 1 (right after equation 1).
2) "Length: the paper and in particular the method section is too long. The method section extends over 8 pages - very detailed, but too detailed for a scientific paper. This should not be a report to an agency. Please constrain the method section much more and refrain from writing the section like a code. Describe what the models does, what physical equation you use, but not how these are taken in the code. If you think this is really necessary, a flow chart with the main model parts would be much more helpful."
   We addressed your critics on "writing the section like a code" and could shorten the methods section for RM2 by 20%. We believe that an even greater shortening would lead to a poorer understanding of the text. Two comprehensive equations could be omitted. We also shortened passages from other sections, skipped repetitions and avoided accentuations. RM2 is about 1000 words (about 10%) shorter than RM1. It is about 21 pages in HESS's final two-column version incl. 6-7 pages of abstract, references, appendix etc.

3) “No selling of the model: there is no need to sell the model to the scientific community. Either it is good, provides new ideas and is superior to other models or it is not. So please refrain from sentences like this: "seems to be the first model since long that ..... Not particularly innovative, but remarkably successful...." This is not a good scientific style.”
   We avoided expressions awaking the impression of selling our work.
4) “I also think that including temperature or other meteorological variables into the model framework may further improve the model. Measurement of precipitation is certainly difficult and hence a model driven by snow depth makes a lot of sense. But how much could the model improve if these observations would be considered. This would be scientifically very interesting. If you can show, that this would not improve the model much (at least showing it at one location), this would certainly strengthen the realist of your study.”
   Including temperature etc. would go directly against what we want to present with the ΔSNOW model for the reasons given in the abstract and introduction. The model is intended to run without any additional meteorological input to make efficient use of long term snow depth series not accompanied by e.g. temperature recordings.
   The new model was intensively compared to other models that need no further meteorological input. Its accuracy was also compared to the respective accuracy measures that are available for more complex models. Running thermodynamic models of several complexities aside the ΔSNOW model and compare their results for SWE would be an interesting experiment for the future. However, it would burst this paper’s limits and would definitely deserve a separate journal article.

Reviewer #1’s letter and our answers and comments.

Dear authors,
the manuscript has profited from the improvements you have realized.
We agree and thank for the elaborate comments.

However, several important issues still deserve attention, some of which have been addressed in the first round already, but they are not sufficiently solved. These issues should be seriously elaborated on prior to publication of the manuscript as full paper.
This is not true. *All* issues have been addressed and seriously elaborated (see author comments of the discussion). The vast majority of the reviewers’ comments led to changes, some issues were defended. “Sufficiently solved” is a very subjective expression (cf. reviewer #2).

(1): it is still not entirely clear what the ultimate aim of your presented work is.
We changed and sharpened the wording of the Introduction, especially of the motivation section 1.3, which outlines the paper’s aim. We also moved a paragraph from section 1.1 to the opening section to provide more of a “set up” for the whole paper.

Main points: (i) You say you „present a new model, a new method to simulate SWE.“
The above mentioned phrase is not used in RM1. We avoided the terms “model” and “snow model” in RM1. Following reviewer comments of the discussion phase we declared, that we present a “new method”. We reconsidered the issue and stepped back for RM1, see below.

Usually a snow modeller would expect a "simulation of SWE" to be based on meteorological forcings (whichever, depending on the empirical or more physical type of the model).
We cannot agree on that. There are snow models based on meteorological forcings, we clearly define

them as “thermodynamic models” (like Langlois et al., 2009), and there are snow models that rely on snow depth and non-meteorological variables like date, altitude etc. The latter we define as “empirical regression models” (ERMs; in the tradition of Avanci et al., 2015) since those are *only* depending on regressions, but do not resolve any thermodynamic processes. The third category that we identify, are the “semi-empirical snow models” (following Gruber, 2014), where the ΔSNOW model belongs to. The latter do not resolve thermodynamic processes, but make use of basic rules of densification. They are not “fully” but “semi”-empirical, since these densification rules are part of theoretical physical concepts like, e.g., Newtonian viscosity. These definitions are very well described in RM1, esp. in Table 1, but we have again sharpened the wording for RM2 and introduced subsections for each model type for better differentiation.
If a “snow modeler” would only expect a meteorological-forcings-based model for a “simulation of SWE”, he/she ignores the whole bunch of ERMs and semi-empirical models, but is only open for thermodynamic models. What else than “simulation of SWE” do all ERMs and semi-empirical models (like references in *) do?

What you present to the reader effectively is a layered snow density evolution parameterization for ephemeral snow packs, from which - of course - swe can be derived. Maybe this could be expressed more precisely.
We identify a misunderstanding: Reviewer #1 seems to only count thermodynamic models as legitimate snow models. ERMs and semi-empirical models are claimed to be “density parametrizations”. This is different to what was done in the past. Many ERMs and semi-empirical models have been presented as “models”. Yes, one could argue that ERMs are only parametrizations of density (or SWE) and do not deserve being called snow models. Still, this is in contrast with authors like Jonas et al. (2009) or McCreight and Small (2014), who naturally call their approaches “models”. But semi-empirical models are definitely more than pure parametrizations since they use, e.g., viscosity physics etc. Why should the ΔSNOW model only be a “layered snow density evolution parameterization”? We already called it (only) a “new method” in RM1 rather than a “new model” as we followed former reviewer comments. In this respect this was more modest than most presentations of ERMs, although our method is more complex.
Having revisited this issue, we now stepped back and again call the ΔSNOW model a “model” rather than a “method” for RM2, since also less complicated approaches have naturally been called models in the past. Maybe it seems offensive to have the word “model” in our initial model’s name (ΔSNOW.MODEL). We now simply call it “ΔSNOW”. (This slightly changes the title of the paper.)

(ii) There is still no strong argument to resign from using available meteorological observations. Even if your goal is to develop an assimilation scheme that uses nothing more than recorded HS, a comparison of this scheme to a model using at least temperature records (generally available, extendable with humidity for phase detection) would be scientifically comprehensible approach.
Isn’t the lack of meteorological observations (aside snow depth) a *very* strong argument? Are all the other papers (*) on ERMs and semi-empirical models no “scientifically comprehensible approaches”?

There should be locations where even precipitation is available, allowing for extending the classical temperature index approach (and, e.g., including rain-on-snow). However, I do agree that this is an option - but a crucial one, as long as you call your model a snow model.

1) “as long as you call your model a snow model”: We actually did not call it a snow model in the RM1, but we now call it a model in RM2 again, since also much simpler approaches have been called models in the past (see above).
2) There are only very few places in the Alps with regular SWE measurements *and* daily HS measurements *and* temperature/precipitation recordings for many years. It was quite hard already to put together those 6 Austrian and 9 Swiss stations with SWE and HS

measurements. We emphasized this even more for RM2 (section 2.3.1). The fact that the Δsnow model requires no other input than snow depth makes it also very worthy to derive SWE from remotely sensed HS recordings. This aspect was also emphasized for RM2.

3) Of course temperature, precipitation etc. could be included in the ΔSNOW model. However, this step would be nothing innovative. Many thermodynamic models can be run with HS, temperature, precipitation (and parametrizations of anything else). First of all: Why should one try to improve a method that should work without a certain input by using this respective input? That makes no sense.

(iii) As the paper is now, it seems that the ultimate goal of your numerical experiment was the snow load map, to be developed without meteorological recordings (because they are not available everywhere for the entire period). If this is the case, the paper should better put this fact in the focus and follow a respective structure and argumentation. If your goal is the model development and comparison with its peers, as you present it in the validation/comparison section, you should focus on the latter and remove much of the rest that does not contribute to this goal. Please make the paper as much as possible clear in its intention.
This criticism was addressed in the revision already. RM1 is about a method to calculate SWE. In the appendix an example of application is given by providing the snow load map.

(2) I still strongly recommend to separate any mathematical formulation of empirical/physical relations from the implementation of the model (code, modules etc.). Swapping two different types of elements in the model along the model description is more puzzling than helpful. Better present the equations of the model in the text of the manuscript, and the code (with its structure – „modules" in your particular implementation) as flow process chart in the appendix.
We made comprehensive changes for RM2 addressing this, but still permit ourselves to provide the mathematical formulations of empirical/physical relations implemented in the model.

(3) The language still is common speech like in many formulations, e.g. „a bit of dexterity“ (39), "with the help of De Michele (2013)“ (212) or „For the ΔSNOW.MODEL this kind of high error tolerance of ρ0 is a rather feeble argument to use a power law“ (202/203). Also better avoid all formulations where the model, the code or a process „does“ or „decides“ or „ignores“ something; It is the modeller who does or decides or ignores. I generally recommend professional native English language support.
We changed accordingly for RM2, and engaged the service of a native speaking professional.

(4) The manuscript should be significantly shortened.
The manuscript was already shortened by 3000 words after the discussion. RM1 is about 24.5 pages in HESS final two-column layout. We could further shorten the text for RM2, mainly in the method section. RM2 should be about 21 pages.

There are entire paragraphs with long explanations that could be replaced by just presenting the mathematical formulas with the respective references (e.g., 384-404).
This paragraph (lines 384-404) is key. How can one calibrate physical parameters without defining their potential limits in advance? Furthermore, these paragraphs are the base of the classification provided in the results section.

And, if you have shaped the focus of the experiments, much of the rest can be omitted. Examples: (i) there is no need to argue the advantages of HS measurements against anything else (chapter 1.1) - long-term observations are always valuable.
Chapter 1.1 is about the big discrepancy between HS and SWE measurements, which is worth being addressed in this context.

(ii) 292-299.
We strongly shortened this paragraph for RM2.

(iii) 349-351: it seems that what you call „scaling" is sort of an assimilation of the HS observations into the current model state by means of an replacement/adjustment of respective variables. Why not say this this with few sentences and skip the rest?
We could only slightly shorten this section. The way how the "scaling" is done is crucial, novel, innovative, not at all straight forward (see App. B), and therefore worth describing.

Other details I came across (exemplary, several occurences of the respective type in the text):
- 89ff: your division of SWE models in the „thermodynamic" and „empirical regression" type of model is unusual (more usual are energy balance vs. temperature-index). Maybe it should clearly be indicated here that you use the expressions for density models?
What we mean with "thermodynamic", "empirical regression" and "semi-empirical" models is clearly defined in the manuscript and supported by references, e.g. Langlois et al. (2009) or Gruber (2014), see above. In many other contexts the differentiation between energy balance and temperature-index models is necessary (In this respect it is "more usual".), but energy balance and temperature-index models are both encompassed by "thermodynamic" models. However, we need the cutting line between "thermodynamic" and "empirical regression" models here.

- entire chapter 1.2: this chapter should lead to the next one, the motivation. However, a clear argumentation (following the Rango/Martinec model description) of what is missing there, or what could/should be improved, is not given. The last sentence seems to loosly tight to what was stated before
We sub-divided the chapter in three parts for RM2. The argumentation is given in the following motivation section (1.3), which was sharpened in this respect. We also shifted the last sentence of 1.2 to 1.3.

- 194: „For non-zero snow depth observations": how many?
We do not understand the question. This is a description of the model concept. It is irrelevant how many non-zero snow depth observations there are. The model works for ephemeral snow packs lasting, e.g., three days as well as for, e.g., high-altitude snow packs lasting many months. We estimate a limit of 200 days in the manuscript.
- 220: „constituted by the sum of loads overlying layer i": something is missing here to make the sentence complete
The sentence is correct. We changed "constituted" to "induced" for RM2.

- 217, 224: avoid „today's …" (better „the most current" or the „actual")
We changed to the suggested version.

- 230: „cannot be observed": why? Did you make an attempt to illustrate density recordings and analyze if and when these reach a maximum value?
No, we did not illustrate density recordings here, but many other did before. There is no "natural constant" that defines a seasonal bulk density maximum, but it is a model constant. That's what we wanted to express here. We removed this phrase.

- 237: better formulate that the age of the layer is a surrogate for the physical processes and their evolution and effect over time
Changed to "Equation 2 links the densification rate to the layer age, but indirectly by the use of density, and not directly as it was the case with Martinec and Rango (1991)'s power law approach.

Consequently, ΔSNOW's compaction is not directly depending on layer age, which is a prerequisite for the functioning of the *Overburden Submodule* (Sect. 2.2.1)"

- 287-290: this certainly depends on the time step of the model
Yes. As this statement is already in brackets we avoid being even more precise here.

- 325: „in the model world": what is a „world" here?
We avoid this for RM2.

- 335: „a small "stretching" of the snowpack is necessary": better avoid such formulations: there is no stretching, but a modification of state variables
The wording is changed for RM2, as other pictorial expressions (see above). Still, "stretching" is presented/defined here as the opposite of compaction. Compaction is (naturally) accepted as an expression, therefore also "stretching" must be acceptable, even more as it is put in quotation marks.

- 387: „Sub-daily means of new snow densities are lower": when and why? Please explain
See respective reference.

- 398: „7.62237 × 106 Pa s": the number of decimal places should be reduced to one
This is a literal quote. It is clearly marked as such and cannot be changed. For the values we found, we use "two significant figures" (except for ρ_{max}). See AC3 of the discussion phase.

- table A2: correct „valibration"
Changed accordingly for RM2.
- etc. ...

* Guyennon et al. (2019), Pistocchi (2016), Gruber (2014), Sturm et al. (2010), Jonas et al. (2009), Mizukami and Perica (2008), McCreight and Small (2014), Martinec (1977), Martinec and Rango (1991), Rohrer and Braun (1994), and many others (see e.g. Avanzi et al., 2015)

Changes from RM1 to RM2 that are not related to referee and editor suggestions:

- Capital letters are omitted in title and section names (following a native's suggestion).
- The abstract was rewritten.
- Citation Paterson (1998) was changed to newer Cuffey and Paterson (2010).
- In section 4.8 an outlook was given citing Lievens et al. (2019).