Response to review:

We would like to thank the reviewer for the constructive comments. We hope that the plots included in this document convince the reviewer that the Richards routine is acting in a numerically stable manner and the non-physical results are caused by the feedbacks discussed in the text. Therefore, we still feel that validation on lysimeter and snow pits data will not add to this work. The feedback issues are similar to the issues that the reviewer faced when developing a similar routine in the SNOWPACK model. However SNOWPACK and Crocus have different structures and slightly different equations (physical or empirical), so the issues presented are similar but unique to Crocus.

The abstract and conclusion have had major changes to highlight that the routine is complete but does not couple well in the Crocus or SURFEX model in the current state. The model code has been revised to ensure a free-flowing bottom boundary and a head based pre-wetting, as recommended by the reviewer. Relevant sections have been revised describing the new pre-wetting mechanism and the free-flowing bottom boundary. Figures 2, 3 and 6-10 have also been reproduced based on the revised model. Similarly, the results section has been revised to reflect the new bottom boundary and pre-wetting.

Response to comments:

For this section the authors' responses are shown in black, where the reviewer's comments are in blue. Quotes ("") and *italic font* show changes to the manuscript.

1 general comments:

The manuscript by d'Amboise et al. discusses the implementation of a solver for Richards equation into the detailed, multi-layer Crocus snow model. This equation describes water flow in porous media, such as snow, and previous studies have already shown that snowpack models can benefit from implementing this equation. It generally seems to provide a better representation of liquid water content (LWC) distributions and snowpack runoff behaviour. In that sense, the study represents an important step for the Crocus snow model. Also the study can be considered an important independent verification of the results achieved with the SNOWPACK model, where the solver for Richards equation has been found to considerably improve the description of liquid water flow in snow in several aspects. I found the manuscript well written and pleasant to read, but there are also some language and grammar issues (see technical corrections below). I value and appreciate the effort undertaken by the Crocus team. Based on my experience, I know that introducing new routines to a snowpack model with such a big impact as a new water percolation routine is a serious and difficult piece of work. My first main concern with the study presented by the authors is that the results presented here are not convincingly showing that the model behaves numerically stable. Distributions of liquid water content look different from distributions achieved with the SNOWPACK model. The absence of a comparison with field data of profiles of liquid water content or snowpack runoff makes it impossible to judge the validity of these results. I will do a few suggestions for additional verification of the numerical scheme, which I hope will provide convincing evidence that the model behaves numerically stable. My second main concern is that the general message of the manuscript is not clear and very open and may potentially confuse readers (see below).

We based our work off the reviewer's description of the routine that was implemented in the SNOWPACK model and described in Wever et al 2014. The review states "the study can be used as an independent verification of the results achieved with the SNOWPACK model". However there are many small differences between the Crocus and SNOWPACK models, specifically in the model structure and some of the physics or empirical relation.

It is important to understand that issues that deal with coupling and feedback are unique to the Crocus model, but probably quite similar to issues faced when implemented in the SNOWPACK model. However, issues with parameter sets that are not a result of feedback will be common for both SNOWPACK and Crocus (grain size of crusts, strong dependence of density and grain size). Therefore, conclusions we draw from the Richards routine inside the Crocus model, may not directly apply to the Richards routine inside the SNOWPACK model.

To address the numerical stability of the routine, we turned off the two largest feedback contributions, the snow metamorphism and compaction routines (SNOWCROMETAMO & SNOWCROCOMPACTN), see Fig. A.

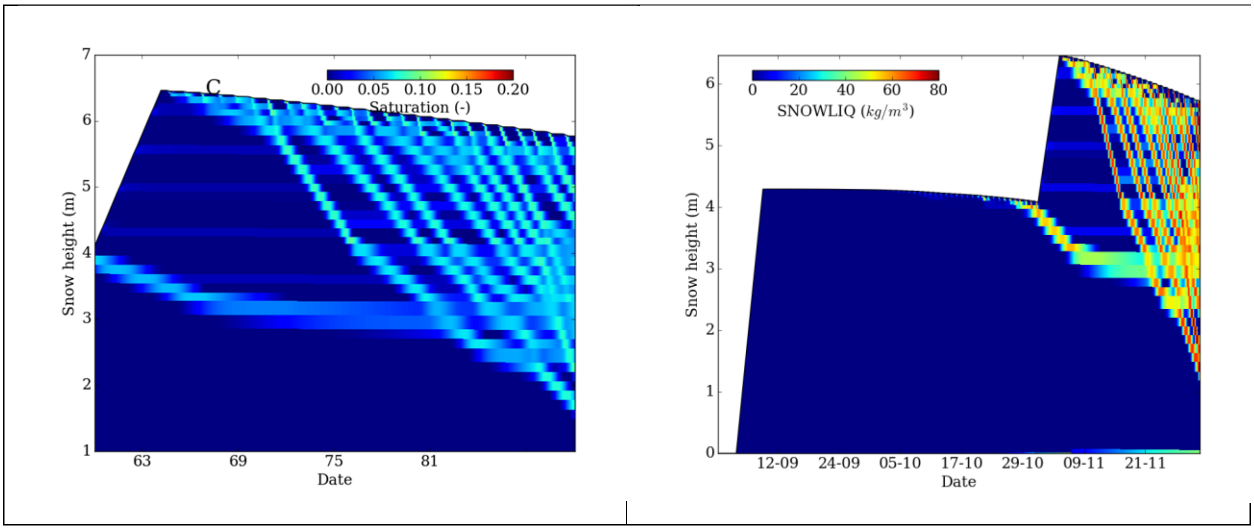

Figure A shows results from the synthetic data set run without the compaction and metamorphism routines. Note that the snow layers are comprised of new snow grains at low density.

Figure A shows the Richards routine behavior is stable, without any “oscillations” or the wet, dry pattern which is seen in Fig 9 & 11 of the original manuscript.

Figure B shows the same data set with the metamorphism routine still turned off, but with the compaction routine turned on.

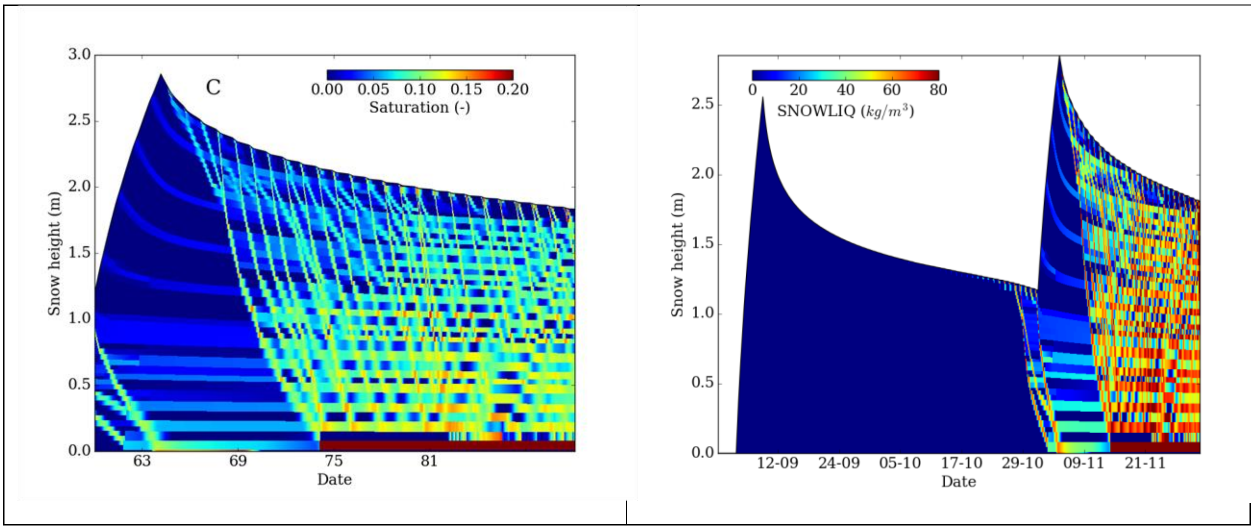

Figure B shows results from the synthetic data set run with the compaction and without the metamorphism routine.

The oscillating pattern appears again. We show here that the pattern of wet, dry layers is not a result of a numerically unstable solution or oscillating solution in the Richards routine, but most likely an effect of an incomplete description of wet snow compaction.

Results (not shown) which used the metamorphism routine and not the compaction routine are more difficult to associate with the striped pattern. Water percolated down only after the LWC became very high > 20%. Water percolated very slow and did not reached the bottom of the snow pack by the end of the synthetic simulation, when run without compaction routine. It is very likely that the high LWCs made by the metamorphism routine will amplify the striped pattern made by the compaction routine.

The second main concern is that the message of the manuscript is not very clear. We agree that the manuscript should be more forward with the message.

We have addressed this issue by adding a few sentences to the abstract and conclusion about steps needed for further development of Crocus and the Richards routine. Together with addressing many of the points made below we feel it is clearer now.

2 Specific comments

I have the following remarks related to the numerical scheme:

2.1

Especially the alternating wet and dry snow layers shown in Fig. 9 and 11, and discussed in p10,l27, are very suspicious. It looks like a numerically oscillating solution. If it is a true LWC distribution, it is recommended to have a higher vertical resolution in the simulations (i.e., more snow layers) in order to better represent the strong gradients between the wet and dry snow layers. But the simulated values of 10%-15% seem unrealistic. Such high values may occur occasionally above capillary barriers, as shown in the work by Avanzi et al. (2016), but I'm not convinced that it should happen so regularly in the simulations shown in Fig. 9 and 11. Particularly because the artificial large snow falls create a very homogeneous stratification, such that ponding is not expected to occur. So actually I wonder if this is not a representation of the fact that the Crank-Nicolson scheme can be prone to spurious oscillations? As far as I know, Crank-Nicolson schemes are generally considered globally stable, but irregular initial conditions may lead to oscillatory behaviour. The current simulations are done with only the optimal time step for Richards equation. But are the model results sensitive to the time step inside the Richards solver? If the model is forced to run with much smaller time steps, are the results different? Or if the model is run with higher grid resolution, or by switching of remeshing, are the results different? As far as I know, the oscillations from the Crank-Nicolson scheme can be reduced by smaller time steps and/or higher grid resolution. Also stability criteria for Crank Nicolson schemes exists, which could be discussed by the authors. Note that I also have a suggestion to initialize the model more stable, see point 6 below, which may also improve numerical stability. Maybe if possible also provide additional motivation for the choice of a semi-implicit scheme instead of a fully implicit one. Is the discretization (Crank-Nicolson + Picard iteration for Richards equation) used here newly developed in this work or has it been applied before? If so, please add the references.
See section 1 (General comments), for remarks on the numerical stability and the root of the "oscillations".

A LWC of 10%- 15% is high but not unrealistic for snow (at least for short periods). However, these are the results obtained with the current state of the Crocus model using the Richards routine. We do not claim that the results of this routine accurately model snow properties, at this point. Yet, the routine has been added to the Crocus model and further developments of the Richards routine is dependent on development of other Crocus routines and/or the parameter sets. The parameter sets have been shown to work with the SNOWPACK model. However we are cautious with them because they make up half of the feedback loop.

Normally the time step is restricted by the Crocus time step and the time step rules given in appendix A1 of the (original and revised) manuscript. To investigate the possible numerical nature of the oscillations, we have restricted the Richards time step permitting maximum values of 60 s, 10 s and 1s. This is based on the notion that numerical stability increases with decreasing time steps, but since the model operates with adjustable time steps, defining the maximum time step mainly affects those situations when the solution procedure behaves well.

Our choice for using Crank-Nicolson with the Picard iteration to solve the Richards equation comes from a previous model developed by L. Oxarango (co-author), based on the h (head) form of the Richards equation (Tinet et al. 2011). There exists an unpublished adaption which solves the mixed form of the Richards equation. The Richards equation has been solved using the Crank-Nicolson and a Picard iteration for a finite element scheme in Paniconi & Putti 1994.

2.2

The mass balance is verified in the Picard scheme with a threshold of 10−4. The authors should add units here, but for now I assume it is the mass balance error in m3 /m3 or kg/m2. I think that this value is set too large to judge mass conserving behaviour of the model. The minimum time step in the solver is 10−10 s (p16,l28). If the solver has a mass balance error of 10−4 with a time step of 10−10 s, this implies a mass balance error of 16 m3 /m−3 /s or 110 kg/m2 /s, or 16 kg/m2 /s, depending on the units. But this is a potentially large mass balance error! This means that if bugs in the numerical scheme or in the implementation of boundary conditions exist, the solver can "cheat" upon the mass balance check by choosing small time steps. Note that in the current version of SNOWPACK, this is also possible. During development of the solver we were particularly paying attention to the smallest time step in combination with the mass balance check. However, we relaxed the condition, by setting a low minimum time step, also allowing the solver to "cheat" the mass balance check. The motivation is to have a more robust solver for end-users. For this review, I analysed the time step distribution for running 15 years of Weissfluhjoch simulations using Richards equation with SNOWPACK, and the smallest time step during this period is about 2x10−5 s. The maximum allowed mass balance error in SNOWPACK is 1x10−10 kg/m2 for the entire model domain. The combination of smallest time step found in this simulation setup, together with the mass balance criterion gives 5x10−6 kg/m2 /s. As can be seen in the Fig. 1 below, these small time steps happen very seldom for 15 years of simulations.

We are calculating mass balance in m because the Crocus variable holding information about wetness (PSNOWLIQ) is in m of water. Units have been added to the manuscript. The conversion to kg/m^2 is achieved by multiplication with the density of water (SI units).

Three mass balances are checked.

1. Full mass balance (mb_15), full snow pack over crocus (15 min) time step.
2. Full mass balance (mb_rch), full snow pack over RCH variable time step.
3. Mesh balance (mesh_bal), layer mass balance over RCH variable time step.

The convergence criteria have been evaluated for mb_rch and mesh_bal.

Figure C (mb_15) and D (time step size). We tested the performance of the models mass balance mb_15 using two limits on the mesh balance, 1E-4 m and 1E-6 m. mb_15 is important to check because it checks the results that are exported to SURFEX/ISBA and other Crocus routines.

Note: The plots have been made with the new implementation of a free flowing bottom boundary and pre-wetting based on head (h) described in sections 2.5 and 2.7 of this response.

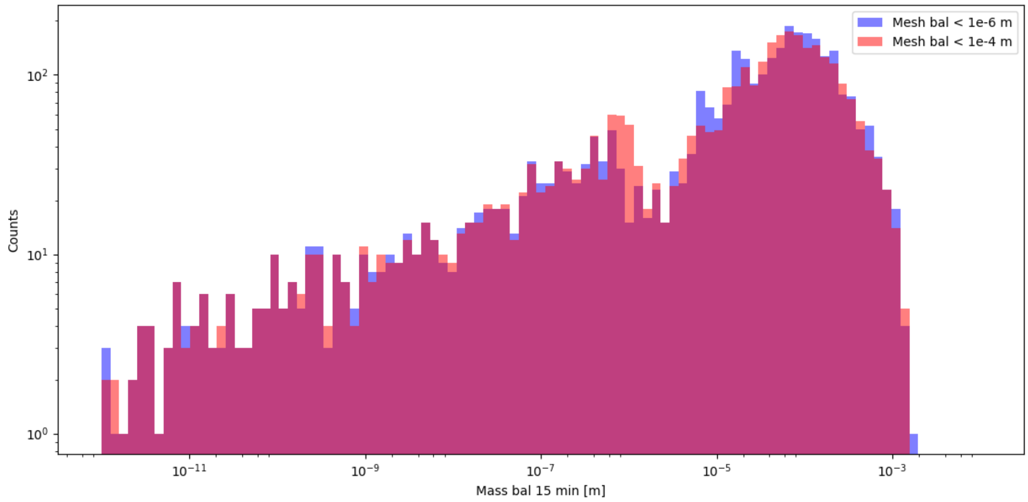

Figure C Difference in the mass balance over 15 minute Crocus time step with mesh balance limits on the Picard iteration. Red =1e-4, blue = 1e-6, pink is both red and blue.

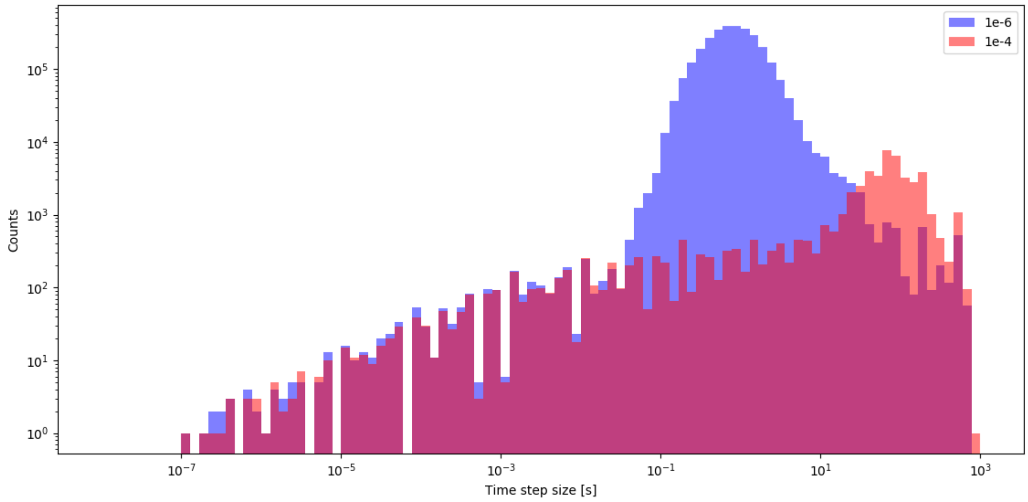

Figure D Difference in the time step duration with mesh balance limits on the Picard iteration. Red =1e-4, blue = 1e-6, pink is both red and blue.

The plots (C and D) show how making a stricter mass balance limit on the Picard iteration loop, has very little effect on the on the mass balance over the 15 minute crocus time step, but has a large effect on the time step size (and therefore model run time). This is because the model "cheats" the mass balance convergence criteria by reducing the time step as the reviewer points out.

The mass balance is one of the many improvements needed for this routine to be used operationally. However, we feel the mass balance is in the appropriate range when compared to the performance of the SNOWPACK model for 3 reasons (see below).

SNOWPACK has a maximum mass balance error rate of 5×10^{-6} $kg/m^2 s$. The rate in Crocus can be estimated from Figure C, and is ~ 1E-3 kg/m^2 s, calculation shown below. There is a difference of 3-order of magnitude between the Crocus and SNOWPACK mass balances for their respective Richards routine.

$$m\,b_{15\mathrm{min}} \sim \frac{0.001[m] * 1000\left[\frac{kg}{m^3}\right]}{900[s]} = 0.00111\left[\frac{kg}{m^2\,s}\right]$$

In the following, we argue why this 3-order-of-magnitude difference is less problematic than it appears at a first look.

1. The Reviewer implies in (section 2.1) that in their simulations SNOWPACK does not often reach LWC in the range of 10-15%. This low water content will cause a low mass balance error which depends on the amount of water involved. In contrast, in our simulations, snow layers are often in this very wet state, hence, the associated mass balance error will be larger. This effect puts the relative difference between SNOWPACK and Crocus into perspective.

2. Our synthetic forcing has been especially designed to challenge the newly implemented Richards routine. There are very strong melt events (during day hours) after the second snow event, yet during night the upper few layers completely freeze. The routine needs to pre-wet these layers and the pressure gradient will be opposing gravity at the transition between wet and dry layers. A natural snow pack will have these conditions a few times over a season, but they are not as common or strong of a melt freeze cycle as seen in the synthetic forcing. We expect that such a demanding situation will be accompanied by a larger mass balance error compared to simulations of more naturally occurring situations

3. Crocus' melt routine sets up strong (and presumably spurious) pressure differences when the Richards routine is entered. The Crocus routine is stepped using constant time steps, during each of which, the Richards routine is entered which applies its own internal time steps. The high pressure gradients are presumably a result of prescribing a water supply mass during the first Richards time step, instead of a constant water supply rate over the entire Crocus time step. This glimpse will be fixed in future work, but is out of the scope for the current study. This would require changes to the melt routines (SNOWCROGONE, SNOWCROLAYERGONE and SNOWCROMELT), most of which would be adapting re-meshing when top layers are melted. Nevertheless, we acknowledge that this shortcoming additionally contributes to the higher mass balance error of our model.

Using a convergence criteria on mesh balance < 1E-6 m or 0.1 kg/m^2. Which sounds like a very relaxed criteria. However, this results in a mb_15 of ~ 0.001 m (during the worst of times) often mb_15 is < 1E-6 m. Despite the three reasons given above there is only a difference of two orders of magnitude on the rate of mass balance error (kg/m^2 s) when compared with the SNOWPACK model.

2.3

A check of the second norm of the deficit vector could help to verify the correct implementation of the matrix inversion to solve the equation. Given: A . b = x, where x is the new solution of the pressure head, then the deficit vector is defined as d = A . b - x, where d is the deficit vector. It should hold that the

second norm of the vector, in case of correct implementation, is (close to) 0. Note that in an attempt to optimize the execution time of the SNOWPACK model, we removed the deficit norm check from the code, after using it first to successfully verify a correct implementation of the solver using this check.
We solve A.x = b using a Thomas algorithm function. The vector d = A.x-b has been calculated with values within the range of 1E-21m to 1E-25 m. The ratio d/b (which normalizes the deficit vector) can be used as an error estimate. We have a range from ~ 1E-15 to 1E-18 (dimensionless) for d/b.

2.4

An overall report for the Crocus model as a whole of the mass balance may also be necessary to verify a correct implementation: Delta SWE = evaporation/condensation + sublimation/deposition + snow/rain fall + runoff?
We show above in section 2.2 of this response, that the mass balance for the Richards routine is sub-par. Therefore we do not feel the mass balance of the full Crocus model is useful at this moment.

2.5

With the numerical scheme for Richards equation in the SNOWPACK model, we found that an improved stability was achieved by initializing dry snow layers (the authors call it "prewetting") based on pressure head instead of LWC. The chosen form of Richards equation uses gradients in pressure head, thus it may be better to reduce gradients in pressure head when initializing the dry snow layers. Therefore, we used the procedure to determine for the whole domain the lowest pressure head in a layer corresponding to a prescribed minimum LWC. This value of the pressure head was then used to initialize dry snow layers, such that only gradients due to gravity are present. It ensures that no snow layer is initialized with a LWC above the prescribed minimum value, while at the same time starting the simulation with a numerically stable pressure head distribution.
This was a very clever hint that is now added to our routine. However the results are not noticeably different in the current state of the routine.

Text added to section 3.1 describing the new pre-wetting method (in the manuscript).

2.6

I noticed in the source code that C (dtheta/dh) is limited to -1−15. For what reason? C is supposed to be the exact derivative of the water retention curve. An artificial cut-off seems unnecessary and may introduce mass balance errors?
This limit has been removed. This limit was never reached during model runs that were used for the manuscript. The limit was put in place to help understanding the evolution of variables and for debugging when we had convergence troubles.

2.7

Section 6.2 and p16,l5-6: why not implement a free-drainage boundary condition at the bottom of the snowpack, instead of all the trouble it seems to give to use the SURFEX upper soil layer? We recently modified the SNOWPACK model such that it can run Richards equation without soil layers, and using a free drainage boundary condition seemed to work well. In SNOWPACK, we implemented freedrainage by setting the flux at the bottom of the lowest snow element similar to the flux at the top of the element, while only allowing downward flux (otherwise setting flux to 0). In case of only one snow element, we set the flux equal to the hydraulic conductivity in this element.
This is a good idea, and we have since added this to the routine with much success. We have previously unsuccessfully tried different bottom boundary conditions excluding interaction with the soil routine.

We added section 3.4.2 to the manuscript dedicated to the free flowing bottom boundary.

3 Manuscript comments

3.1

The message of the manuscript is ultimately very unclear and open. The authors apparently don't trust the new water percolation scheme enough to use it for validation (p16,l9). That basically indicates to readers that this publication is not intended to encourage users of the Crocus model to use the new percolation scheme. But what is then the main message of the manuscript? What are the next steps to improve the trust in the validity of the new routine? Are there any further developments needed or planned? The authors should also provide clear instructions of how to repeat the experiments. I was able to download the source code, but did not find any manual or readme to compile nor did it seem to contain the necessary files to run the test cases.
We have added this link to section 8 "Code and data availability" that is a walk through for running Surfex.Crocus.

https://opensource.umr-cnrm.fr/projects/snowtools/wiki/Basic_functioning_of_SURFEX_without_the_s2m_tool_from_snowtools

We do not push for the routine to be used, but rather point out where further development is needed for this routine to work well inside the SURFEX/Crocus framework. The major developments needed for this routine to be viable are outside the Richards routine, inside other Crocus routines and/or in better microstructure understanding/parameter sets. Therefore we feel that this is a complete work and should be reported on. This manuscript has two main messages:

1. The description of the new routine.
2. Highlighting parts of routines in Crocus that need adaptions/further development in order for the Richards routine to perform well in Crocus.

We note that the second point needs to be stated more clearly in the manuscript, and acted accordingly. Both the abstract and conclusion have significant changes.

3.2

I don't agree with the last sentence of the abstract. First, the absence of validation limits the value of any comment about applicability, but basically the same uncertainties in water retention curves for different snow types and for high density crusts, and also many of the feedback mechanisms are present in the SNOWPACK model too. Nevertheless, we have now demonstrated several times that solving Richards Equation is having usefull applications for the SNOWPACK model, in spite of all the uncertainties. For example for assessing wet snow stability (Wever et al., 2016a; Vera Valero et al., 2016) as well as in a detailed analysis of rain-on-snow events (Würzer et al., 2017) and for reproducing ice layers (Wever et al., 2016b). As shown in Table 1 in Wever et al. (2015), different water retention curves or different methods to determine the hydraulic conductivity at the interface nodes (arithmetic vs geometric) have limited influence on the statistics for runoff, whereas the statistics clearly improve over the bucket scheme. Therefore, I don't agree with the statement of "limited applicability" with the reasons provided in the rest of the sentence.

We can agree that Crocus and SNOWPACK are similar models that have some minor differences, however I think the minor differences can have major influence on the feedback systems in place. Both models "wet snow processes" were developed to be used with the bucket model. The calculations of "bucket size" differs between the Crocus (Vionnet et al 2012) and SNOWPACK (Wever et al. 2014) models. The bucket model for Crocus has a limit defined as a % of pore spaces (set to 5% as default, but may be changed if desired), holding capacity has a linear relationship to density. The SNOWPACK bucket size is a piecewise linear function of pore space. Essentially, Crocus has an almost binary snow wetness model with the bucket model, the SNOWPACKs bucket model seems more dynamic (in relation to pore space). This provokes the idea that wet snow processes are handled differently in SNOWPACK and Crocus.

The feedback is suspected primarily to be between the compaction & metamorphism routines and the parameterizations of the water retention curve and conductivity function. The parameterizations seem to work well inside the SNOWPACK model, despite the theoretical weaknesses that exist with them. However, we were not able to distinguish (for Crocus) if the errors come from the parameterizations (water retention curve, conductivity), other Crocus routines (compaction, metamorphism), or how they interact with each other. Since there are some theoretical problems with the parameterizations and they are the parts of the Richards routine which are part of the feedback mechanism, we cannot exclude the possibility that the parameterization used is the dominant source of error.

To address the issue, a study on microstructure is needed, either expanding the data set that the parameter sets are based on, or investigating the rates of wet snow metamorphism and compaction. Both of these experiments are outside the scope of this study.

We do not feel that our model can reproduce the results that were demonstrated in several recent publications using the SNOWPACK model applying the Richards equation. Therefore we feel the Richards routine as implemented in the SURFEX/Crocus framework has limited applicability in its current state.

The sentence in question is: "We show that the new routine has been implemented in the Crocus model, but due to amplification of parameter uncertainties through a number of feedbacks, meaningful applicability is limited until new and better parameterizations of water retention is developed for different snow types."

We do feel that this sentence should not single out the water retention curve and was changed, see below.

"*We show that the new routine has been implemented in the Crocus model, but due to feedback amplification and parameter uncertainties meaningful applicability is limited*."

3.3

The discussion section is nicely written and provides an interesting introspective discussion about the uncertainties and potential feedback mechanisms in water percolation modelling. Note that, however, many of the feedbacks are hypothesized or based on results of other studies. The manuscript itself does not present material supporting or quantifying the strength of those feedbacks (no validation or sensitivity study). It would be good if the discussion could be made stronger. For example, the authors may want to discuss how water retention curves for crusts potentially look like, and how strong this influences LWC distributions or snowpack runoff?

The Crocus model is not in a state where it works well with the Richards routine. However, we discuss this somewhat unexpected behavior and pinpoint the major sources of uncertainty. We demonstrate that due to the discussed feedbacks, even small structural differences (as for instance between Crocus and SNOWPACK) can be amplified leading to very different results. Furthermore, we outline potential future improvements to avoid these problems. Altogether, we think this is a contribution worthy being reported to enable future scientific progress. However conclusions can still be drawn and supported from the work presented. Much of the discussion pertains to reinterpreting (other researcher's) experimental studies in the context of modeling, which was not done in many of the original experimental studies.

See section 3.10 (of this document) for discussion about crusts layers.

3.4

p4, section 2.1: Maybe explicitly state that the working of the bucket scheme in Crocus is very similar to the SNOWPACK model. p4,l9: Is the bucket size always fixed to 5% of pore space? The sentence following this sentence is a bit confusing, as if there are additional constraints. Should the sentence "For Crocus, ..." not better read "This makes the holding capacity proportional to the density of the snow layer, *but* independent of snow grain type or surrounding environment." It it also not clear what is meant by "surrounding environment" and how it could potentially influence the holding capacity?
There are no additional constraints on the "bucket size". This is stated because it is important that the reader understands that the bucket routine does not account for grain type, or the surrounding environment (layers above or below layer in question, soils water condition etc.). The Richards routine accounts for both of these (unless free flowing boundary at bottom of snow pack is used).

We have added that 5% is a default value and gave example of surrounding environment.

"The "bucket size" or holding capacity has been defined as 5% of a layers pore spaces as default, however this can be adapted if needed. (Vionnet et al., 2012). For Crocus, the holding capacity is proportional to the density of the snow layer but is independent of snow grain type or surrounding environment (adjacent snow layers, soil, ect)."

3.5

p8,l14: please provide a bit more detail on how the amount of evaporation is determined. Atmospheric forcing only provides the latent heat flux, which needs to be partitioned in evaporation and sublimation. How is Crocus doing it? Note that a reason for numerical problems with Richards equation can be when the prescribe evaporation flux exceeds the available water. For this reason SNOWPACK employs a system where the evaporation cannot exceed the amount of water available in the upper element plus the amount of water that can be advected from below given the hydraulic conductivity there. Similar for influx, although typically unrealistic large rainfall rates (>> 200 mm/hr) are necessary to exceed the absorption limit in snow of liquid water. In reality melt ponds form in snow only when liquid water cannot leave the snowpack below, not because the water input rate exceeds the snow absorption capacity.
The evaporation is described in Vionnet et al 2012 section 3.7 "Surface fluxes and surface energy balance". The evaporation is calculated before the percolation routine with the use of two routines SNOWCROBUD and SNOWCROFLUX. Vaporization/condensation can occur when the top snow layer is wet, otherwise mass flux is in a solid state (deposition/sublimation). We did not find problems with fluxes being too high. Evaporation should not affect the Richards routine in the current set up of the model, and rain rate never reaches a value close to 200 mm/ hr.

3.6

p13,l20: "such that the criteria to enter the percolation routine has been met." This is very confusing at it is for the first time mentioned that there are criteria whether or not to enter the percolation routine. Which criteria are meant here?
We added a sentence describing the criteria in Sec. 3.3 in the revised manuscript (Implementation of Richards routine in Crocus).

There are 2 criteria,

1. Number snow layers >= 3, currently bucket routine is used for melt out. (we do not reach melt out in this study)
2. The snowpack has to contain > pre-wetting amount of water, or have a rain rate of > 10^{-10}

"To reduce unnecessary computations the following 2 conditions need to be satisfied before entering the Richards routine:

1. *There must be a snowpack. If there is < 3 layers of snow the bucket model will be used for water percolation.*
2. *There must be liquid water. If $\theta < \theta_{min}$, and the rain flux (over the time step of 15 min) < 10^{-10} m.*

If one of these two conditions are not satisfied Crocus will be run without calling the Richards routine"

3.7

p14,l13: Many examples can be found to show that preferential flow has a much smaller typical spatial scale. See for example Fig. 2 in Techel and Pielmeier (2011), or Fig. 1a in Würzer et al. (2017). Many other examples can be found in literature.

The purpose of this sentence was to show that there is spatial variability that will affect water flow on a scale >1m. Therefore we have taken out the reference for preferential flow and adapted the sentence.

"Hydraulic conductivity has been determined by Calonne et al., (2012) on small snow samples therefore it is possible that the measurements do not represent conductivity at a higher spatial scale because processes that affect the density distribution, grain metamorphism and pore space differ on a meter to several meter scale (Birkeland et al., 1995)."

3.8

p15,l2-3: "this claim needs validation": I think it depends on the application. During the first wetting, grain shape will probably play a very important role. New snow getting wet probably retains much more liquid water initially than the water retention curve developed for melt forms will provide. For wet snow avalanche prediction, the first-wetting is often considered of crucial importance and I think improvements in the description of water flow in new snow and faceted snow (generally less shear strength) are required. On the other hand, for many hydrological applications, often the runoff behaviour during a melt season is important, for which the assumption of melt forms is justified.

The sentence in question is P15 l1 (original manuscript): "This negative feedback on snow grain type could mean the melt forms are the only crystals that need a modeled water retention curve, but this claim needs validation"

This is a matter of wet snow metamorphism speeds, which should be validated experimentally (via microstructure experiment, or a model sensitivity study). We agree that studies using a time scale of "melt season" melt forms should be sufficient. However if the rate is very fast (~30 min), melt forms may be justified for applications which use a shorter time scale.

The reviewer brings up a good point here. The bucket model has been shown to provide reliable results for runoff behavior for longer periods (Brun et al 1989 (fig 9), Brun et al 1992, (fig, 6)). Motivation to add the Richards routine to Crocus come from a desire to better simulate the first wetting of the snowpack, and the distribution of water over the snow layers. Improvements on longer scale application are welcome, but a long term goal is to develop snowpack models to a point where first wetting can be modeled. The work done for this manuscript, does not reach this point but it is a necessary first step, from switching from the empirical routines to physical ones.

3.9

p15,l19: This is a bit confusing wording, as principally, I would say that snow layers are initialized with the "pre-wetting" amount. But here, it is probably meant initialization when the routine is being called during a Crocus time-step

P 15 l19 "However, 83% of snow layers were initialized in the Richards routine with a LWC< 10%"

We changed it to "*…83% of the snow layers entered the Richards routine…*"

3.10

p15,l26-28: Although I agree with the statement, it cannot be considered a conclusion of *this* work. It has not been demonstrated that the water flux over the crust is over- or underestimated (no validation done), neither has it been shown that the simulations are sensitive to the hydraulic parameters for crusts (no sensitivity study done)

p2, l17 we state results from Jordan (1995), that shows that waters behavior around crust layers is not well understood, some crusts promote flow and others prevent it. Therefore, the water retention curve has to have both high and low suctions when compared to surrounding snow layers. Regardless of waters behavior we are certain that the grain size used in the parameterization of the retention curve and the conductivity function is wrong, since crusts don't have individual snow grains. Because theory breaks down when it comes to crust layers there is no amount of quantifiable validation that can justify that crust are modeled in a physical way with the current parameterization. If the model is not sensitive to this error is a different question that should be validated/quantified. As far as we know this concern has not been addressed in other studies.

Our justification for using this parameterizations in the Richards routine is that there are no other parameterizations for ice crusts available.

3.11

Fig. 5: This figure is only mentioned once in the manuscript, and is not discussed at all. Please discuss the agreement between model and observation, or remove the snow profile.

We removed the figure.

3.12

I think the manuscript should not only show results for LWC distributions inside the snowpack, but also snowpack runoff

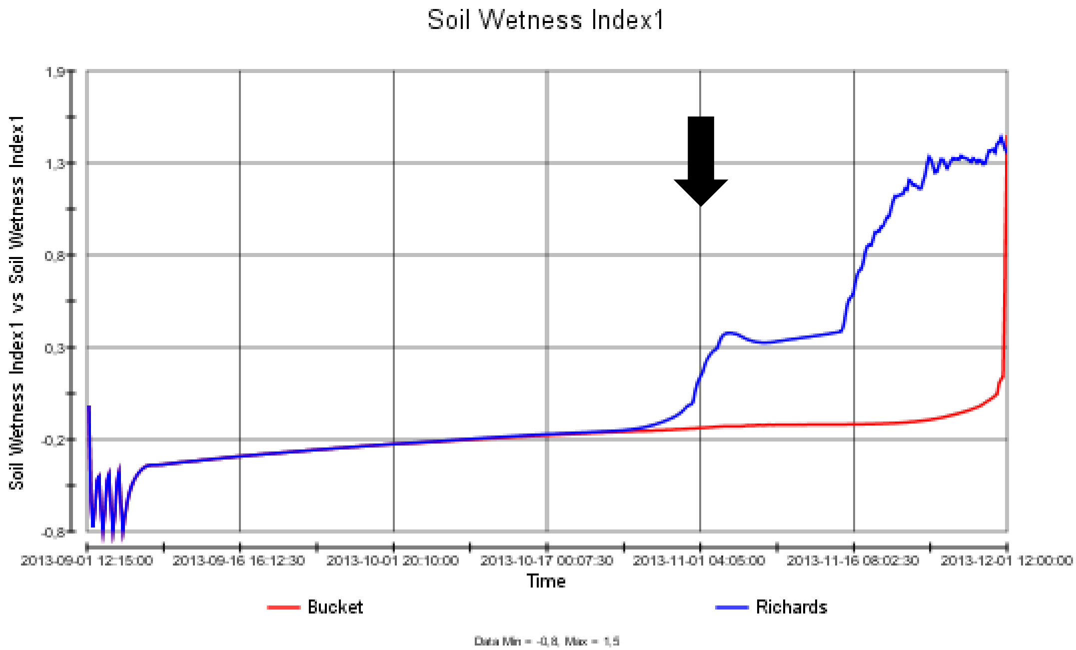

Figure C Top soil layers water content for the synthetic data set, red is bucket routine, blue is Richards routine. Black arrow shows timing of 2nd snow event.

Figure C shows a plot of the top soil layers LWC, which is not snowpack runoff but a result of snowpack runoff. The synthetic data set was designed to highlight differences between the routines over a wide range of water input rates (melt in this case). Therefore, it is not expected that the bucket routine and the Richards routine have similar flux to the soil with the synthetic data set.

This plot shows that the Richards routine is able to move water from the snowpack to the soil column. The timing of the soil wetness increase agrees well with the timing of water percolation in Fig. 9A (black arrow shows time of 2nd snow event).

3 Technical corrections / minor comments:

• General: note that Calonne et al. (2012) determined permeability, which can be related to conductivity. But in principle they did not publish conductivity experiments.

We have corrected this in the revised manuscript.

• General: please mention somewhere the CPU time needed to run the simulations, compared to the bucket scheme.

There are many factors that affect the run time for simulations. Using our "default" settings for this manuscript:

Convergence
head error < 1e-3 m
theta error < 1e-5 unitless
mesh bal <1e-4 m
Crocus time step of 900 s
*Output saved 900 s

* Normal Crocus default saves output every 3 hr. or 10800 s. With Richards routine the plots look like they have randomly high pixels when the plot resolution is 3 hr. Saving the output so frequently is a major contributor to the run time.

The synthetic simulation with Richards took 9m 38s where with the bucket it took 8m 46s. Adapting the pre-wetting amount, convergences criteria (mesh_bal limit), time step and saving output resolution has a major influence over the runtime. We think it is not useful and misleading to report on the runtime of Crocus in this manuscript because many of the variables will likely change with further development of the Crocus model.

• General: there is a change of tense sometimes, compare for example section 4.3 with 3.4.

We have corrected this in the revised manuscript.

• General: sometimes "Figure" and sometimes "Fig." is used to refer to figures.

We follow the rule of GMD, Fig. should be used unless at the start of a sentence where Figure is used.

• Abstract: "this routine is based on". Why the wording "based on"? I would write here: "this routine solves Richards equation".

We have changed this in the revised manuscript.

• p2,l2-3: note that simulations using Richards Equation have already been used for the assessment of wet snow avalanche activity (Wever et al., 2016a) and for determining the initial conditions for wet snow avalanche dynamics simulations (Vera Valero et al., 2016). Given that the authors discuss this topic in particular, they may consider citing these studies here.

Below sentence with citations have been added.

“A physically based water percolation model has been used along with weather station data to improve forecasts of wet snow stability (Wever et al., 2016a), and for determining initial conditions for simulations of avalanche dynamics (Vera Valero et al., 2016).”

• p2,l16: This would not be the way I would explain the precondition for flow fingering, but I also have not the evidence to object against it. Maybe verify with DiCarlo (2013)? I think this is a more up-to-date citation that may be cited here aswell.

The sentence in question (below) was not changed in the modified manuscript.

“Pressure gradients acting against the gravity may induce the formation of preferential flow channels, as flow channels in soil occur where capillary pressure gradient opposes the waters flow direction (Philip, 1975).”

The reference suggested DiCarlo (2013) is a detailed review of finger flow in sand and soil with a strong focus on characterizing finger flow behavior for modeling. Since the Richards routine does not account for finger flow we do not think it is necessary to described flow fingers in such detail. The sentences purpose is to state that finger flow occurs and probably has a connection with suction (were pressure gradients oppose gravity).

We believe this is similar to the criteria used by Würzer et al. (2017) for initiating finger flow in the snowpack. However, we are unsure if the term “pressure head” in Würzer et al. (2017) refers to, pressure head or suction (negative pressure head).

• p3,l3: "Greenland Ice Sheet" (capitalized)

We have corrected this in the revised manuscript.

• p3,l8: "ice crusts": in line with the International Classification (Fierz et al., 2009), this should be "layers". The mentioned study concerned more with ice layers than with ice crusts. Similar p3,l4: lenses are discontinuous ice layers. In this case, I think it should be layers rather than lenses.

We have corrected this in the revised manuscript.

• p6,l19: "kr" should be kr.

We have corrected this in the revised manuscript.

• p7,l17: Here and elsewhere: I prefer "conductivity of snow"

We have changed this in the revised manuscript.

• p7,eq 13: I assume the second equation should read Δzbot

We have corrected this in the revised manuscript.

• p7,l14: I assume the reference should point to Eq. 13 instead of 11.

We have corrected this in the revised manuscript.

• p7,l24: citation style of Celia et al. is wrong (without parenthesis)

We have corrected this in the revised manuscript.

• p8,l3: "computation step" is a vague term. I think this refers to the iteration level k+1? Maybe write: "the pressure head h at iteration level k+1.".

We have corrected this in the revised manuscript.

• p8,l29: maybe specify: "Air temperature become as cold ..."

We have corrected this in the revised manuscript.

• p9,l14-15: please rewrite sentence

We have corrected this in the revised manuscript.

• p10,l2: wrong figure reference

We have corrected this in the revised manuscript.

• p11,l10: "there is a complicated one-to-many relationship ..." Actually it seems to be very simple: below 2x10−5 there seems to be almost no effect, so the prewetting should just be below this value...

This figure and sentence have been taken out of the revised manuscript.

• p11,l14-15: I understand what is meant here, but it may be unclear for readers without a strong snow modelling background. I would explain the remeshing procedure in the model description.

"The alternating dry wet pattern appears before the second snow event when T_{Crocus} = 60 sec. The pattern smooths out during the second snow event, probably due to rearranging of the snow layer sizes."

These two sentence have been removed because the figure no longer shows this effect due to the head based pre-wetting and the free flowing boundary condition.

• p12,l10: typo: "witch"

We have corrected this in the revised manuscript.

• p12,l21: "simulated data sets simulation" I suggest "synthetic data sets simulations"

This sentence has been removed, because results show free-flowing bottom boundary.

• p13,l17: "and but does not have" please reformulate

This whole section has been changed.

• Appendix A: This time stepping is method is very similar to the one I used, and I based it on the work by Paniconi and Putti (1994). Maybe give them credits by citing their work?

We added this citation.

• p14,l19: "on visual grain size measurements"

We have corrected this in the revised manuscript.

• p15,l7: "where, " —> ", where"

We have corrected this in the revised manuscript.

• p15,l11: "grain" —> "grains"

We have corrected this in the revised manuscript.

• p16,l9: "is used to deal"

We have corrected this in the revised manuscript.

• p16,l24: "criteria are met"

We have corrected this in the revised manuscript.

• p16,l28: "within"

We have corrected this in the revised manuscript.

• p17: eq. A1 is not numbered as such
We have corrected this in the revised manuscript.
• p17,l13: "lower density snow that found": please reformulate.
Changed to "*For use in the Richards routine this parameter set would have to be applied to lower density snow than used to create it.*"
• p17,l17: "density not included"
We have corrected this in the revised manuscript.
• p17,l19-20: please reformulate. This sentence cannot start with "while".
While has been taken out.
• Appendix B.3: Note that it should read "Daanen" and not "Dannen".
We have corrected this in the revised manuscript.
• References: a few still point to discussion papers, where final papers have already been accepted and published, for example: Avanzi et al. (2015) and Wever et al. (2016). Please provide DOIs consistently when available.
References have been updated as needed.
• Fig. 5: Specify here also from which date the snow profile is.
Figure removed. See section 3.11 of this document.
• Fig. 7: subfigure B is wrongly labelled C
We have remade most of the figures and made sure they had correct titles and labels.

Brun, E., Martin, E., Simon, V., Gendre, C. and Coleou, C.: An Energy and Mass Model of Snow Cover Suitable for Operational Avalanche Forecasting, J. Glaciol., 35(121), 333–342, doi:10.3198/1989JoG35-121-333-342, 1989.

Brun, E., P. David, M. Sudul, and G. Brunot. "A Numerical Model to Simulate Snow-Cover Stratigraphy for Operational Avalanche Forecasting." Journal of Glaciology 38, no. 128 (January 1, 1992): 13–22. doi:10.3198/1992JoG38-128-13-22.

Birkeland, K. W., K. J. Hansen, and R. L. Brown. "The Spatial Variability of Snow Resistance on Potential Avalanche Slopes." Journal of Glaciology 41, no. 137 (1995): 183–190.

Jordan, R.: Effects of Capillary Discontinuities on Water Flow Retention in Layered Snow covers, Def. Sci. J., 45(2), 79, 1995.

Paniconi, Claudio, and Mario Putti. "A Comparison of Picard and Newton Iteration in the Numerical Solution of Multidimensional Variably Saturated Flow Problems." *Water Resources Research* 30, no. 12 (1994): 3357–3374.

Tinet, A. J., L. Oxarango, R. Bayard, H. Benbelkacem, Guillaume Stoltz, M. J. Staub, and J.-P. Gourc. "Experimental and Theoretical Assessment of the Multi-Domain Flow Behaviour in a Waste Body during Leachate Infiltration." *Waste Management* 31, no. 8 (2011): 1797–1806.

Vionnet, V., Brun, E., Morin, S., Boone, A., Faroux, S., Le Moigne, P., Martin, E. and Willemet, J.-M.: The detailed snowpack scheme Crocus and its implementation in SURFEX v7.2, Geosci Model Dev, 5(3), 773–791, doi:10.5194/gmd-5-773-2012, 2012.

Würzer, Sebastian, Wever, Nander, Juras, Roman, Lehning, Roman, and Jonas, Tobias, "Modelling Liquid Water Transport in Snow under Rain-on-Snow Conditions-Considering Preferential Flow." Hydrology and Earth System Sciences 21, no. 3: 1741, 2017

Wever, N., Fierz, C., Mitterer, C., Hirashima, H. and Lehning, M.: Solving Richards Equation for snow improves snowpack meltwater runoff estimations in detailed multi-layer snowpack model, The Cryosphere, 8(1), 257–274, doi:10.5194/tc-8-257-2014, 2014.