We are thankfull Richard Essery for his constructive review. Please find below our point by point response to the review. The comment of the referee are shown in blue and our response in black below. Proposed modifications of the manuscript are shown in green with page and line numbering corresponding to the preprint version of the article.

I greatly enjoyed reading this paper. The method for efficiently coupling the nonlinear surface energy balance to the linear subsurface heat conduction is a clever piece of matrix algebra, but it is not just that; it directly relates to a point of contention in the lively interactive discussions of Brun et al. (2022) and Potocki et al. (2022) concerning the mass balance of the Everest South Col Glacier.

There are limitations, however. Many processes are generally handled sequentially in snow models (Clark et al. 2015), but this paper only couples two of them. Only idealized test cases are shown and not full model performance in real applications.

We agree with the reviewer that two of the limitations of this work (which were also pointed out by other reviewers and the editor) are that: (i) we only tightly-couple two processes (the SEB and the internal heat equation) and leave others (such phase changes or liquid water percolation) sequentially treated and (ii) that we only treat idealized test cases. We were also aware of this potential limitation when doing this study, and wondered if more realistic cases should be analyzed. We eventually decided to leave them out.

Our motivation behind this choice is to allow to focus on a single topic, namely the efficient numerical coupling of the SEB to the internal heat budget in a FVM framework. We worry that introducing other tight-couplings (such as phase changes while solving the heat equation) might make the role of coupling the surface and internal energy budgets less clear, and thus renders this point less-readily available for current FVM models, such as Crocus or COSIPY. Likewise, we decided to focus on test cases without comparisons to direct observations, as it is not be possible to decipher errors due to the numerical implementations (which is the focus of our paper) from errors due to the assumed physics, parametrizations, and forcings (which we do not and cannot not address in this study). Therefore, we think that to meaningfully analyze numerical implementations in terms of cost, accuracy and robustness, the use of simplified test cases is appropriate. We however agree that the test cases should not be too unrealistic if we want their results to be informative of how a numerical scheme might behave in a realistic settings. That it is why we have used realistic forcings and initial conditions.

We now specify our intention more clearly in the text, and clearly explain that our simplifying assumptions are meant to ease the comparison of the numerical implementations of the surface-internal energy budgets, but that our toy-model should not be viewed as proper a snowpack/glacier model as many important components are lacking.

P12 - L330
"Two simple examples, showcasing the differences between numerical treatments, are presented below. We note that these simulations cannot be considered as fully realistic simulations of a snowpack or glacier surface, as many processes, such as the deposition of atmospheric precipitation (rain or snow) or mechanical settling, are lacking. The goal is rather to provide a simplified setting in which the impact of the numerical implementation of the SEB can be analyzed. In the same idea, we do not attempt to compare the simulation results to field observations. Indeed, it would not be possible to decipher errors due to the numerical discretization (the focus of this paper) from errors due to the assumed physics, parametrizations and atmospheric forcings. Nonetheless, in order for

the results to still be informative of how a given numerical implementation might behave in a more realistic setting, we use realistic atmospheric forcing, initial conditions, and physical parametrizations. The first simulation is meant to highlight the behavior of the numerical models when simulating the SEB on a snow-free glacier. The second one focuses on the impact of the model implementations on the simulation of the energy budget of a seasonal snowpack, during the melting period."

From the test case results, I could take contrary conclusions that the added complexity of coupling is not needed and the standard skin-layer formulation is fine as long the time step is not made too large, which is well known ("not too large" could still be prohibitively small for thin layers, though).

Indeed, very reasonable results can be obtained with the standard skin-layer formulation as long as the time step is kept short enough to avoid instabilities. The same conclusion could be made for the Class 1 model (no surface), as long as the top cell is kept thin enough. As all models solve the same equations, they converge to the "true" solution when the spatial and temporal resolution are refined, and the tightly-coupled approach is not expected to yield a different solution.

However, we believe the property of the tightly-coupled approach to accept both large time steps and mesh sizes, while keeping a similar numerical cost, motivates it use over the standard approaches. A numerical stability analysis is provided at this end of this response that shows that the coupled-surface scheme is unconditionally stable, contrary to the standard-skin layer formulation, and thus does not to require the implementation of an adaptative time step strategy. This is now discussed more clearly in the revised manuscript and the numerical stability analysis provided in an appendix.

P21 - L503
"The unstable nature of class 2 models can be shown with a linear stability analysis, provided in Appendix E. Such analysis shows that class 2 models are only conditionally stable, and confirm that instabilities are favored in the case of large time steps and small mesh sizes. We stress that these oscillations can appear even if the time integration of the internal energy budget relies on the Backward Euler method, known for its robustness against instabilities (Fazio, 2001, Butcher, 2008). Our understanding is that the sequential treatment of the standard skin-layer formulation breaks the implicit nature of the time integration by using "lagged" (in other words, explicit rather than implicit) terms. This, combined with the fact that the surface layer does not possess any thermal inertia and that its temperature can thus vary rapidly in time, permits large temperature swings if the time step is too large or the mesh size too small. On the other hand, it can be shown that the two schemes with a tightly-coupled SEB are unconditionally stable (Appendix E), in agreement with the absence of oscillations in their simulations. Notably, the unconditional stability of the coupled-surface scheme proposed in this article entails that the model does not need an adaptive time step size strategy depending on the mesh size. This ensures that it remains robust, regardless of the time step and mesh size.

P26 - L563
"Moreover, a tightly-coupled treatment of the SEB allows unconditional stability, while the standard skin-layer formulation can be unstable and displays large spurious oscillations with large time steps and small mesh sizes."

Specific comments:

Author list
“Brun Fanny” might like to have her name turned around.
We put Fanny's name in the good order.

Introduction

I don't recommend writing a comprehensive review of SEB formulations, but only giving recent examples of applications of a skin layer and no examples using a finite surface layer in the introduction, rather than original model development papers, gives a distorted view. An uncoupled skin layer has been in use for snow models at least as far back as Yamazaki and Kondo (1990). There is a snow surface layer temperature in Anderson (1968).
We added older model development papers in the revised manuscript. Notably we now provide references when discussing finite-top-layer models.

P1 - L20
"To reach this goal, the representation and evolution of the thermodynamical state (that is to say temperature profiles and phase changes) of snowpacks and glaciers are implemented in most numerical snowpack/glacier models (e.g. Anderson, 1976, Brun et al. 1989, Jordan, 1991, Bartelt and Lehning 2002, Liston and Elder, 2006, Vionnet et al. 2012, Sauter et al., 2020)."

P2 - L40
"On the other hand, some FVM implementations do not define a specific temperature associated with the surface, but rather use the temperature of the top-most numerical layer of the domain (i.e. the top layer of the simulated snowpack/glacier) for solving the SEB (Anderson, 1976, Brun et al., 1989, Jordan, 1991, Vionnet et al., 2012, van Kampenhout et al., 2017)."

22- There are many "numerical models" that are not snowpack/glacier models.

We reformulated the sentence to clearly state that by "most numerical models", we want to refer to snowpack/glacier models.

P1 - L20
"To reach this goal, the representation and evolution of the thermodynamical state (that is to say temperature profiles and phase changes) of snowpacks and glaciers are implemented in most numerical snowpack/glacier models "

32- The surface energy balance is described as "profoundly non-linear". Actually, this is a pretty benign nonlinearity in the field of nonlinear equations; it does not have multiple or chaotic solutions.
We removed the word "profoundly" to only state the problem is non-linear, and hence might requires some iterations for the proper solution to be computed.

49- The "infinitely small horizontal layer" would be better described as infinitely thin.
We replaced "small" with "thin".

While Eq. (1) is more generally applicable, it could already be emphasized that this is invariably implemented as a 1D model with T a function of z.

Consistently, with the remark of Reviewer 3, we now state that while the equation remains valid in 3D, we use it in a 1D set-up only as transitionally done in snowpack/surface-glaciers models.

P3 - L70

"In this article, we assume that the snowpack/glacier can be represented as 1D column, and therefore Eq. (1) should be understood as 1D equation."

I think that there will be very few exceptions to this "usually" of allowing snow temperature to exceed the fusion point before calculating melt, but there are examples of models with phase changes over a temperature range in Albert (1983) and Dutra et al. (2010).

As also pointed out by the review of Michael Lehning, several strategies have been proposed to handle phase change in snowpack/glaciers models. We modified the revised manuscript to clearly state that we rely on the method of exceeding the fusion point and then restoring thermodynamic equilibrium as it employed in the majority of snowpack/glacier models, but that alternatives exist. We now also stress that this method of "overshooting" is a form a sequential treatment, to which better treatments have been proposed in the recent literature. Building on this idea, we are currently working on the efficient tightly-coupled resolution of all internal thermodynamic processes, and will address it in a future work.

P3 - L82

"In this article, we follow this simple scheme as it is commonly employed in snowpack and glacier models. That being said, other, more complex, strategies have been proposed in the literature. This notably includes the use of a finite temperature-range over which melt/freezing occurs (e.g. Albert, 1983, Dutra et al., 2010), including melt/refreeze as an additional energy source term (e.g. Bartelt and Lehning, 2002, Wever et al., 2020), or the use of enthalpy as the prognostic variable (e.g. Meyer and Hewitt, 2017, Tubini et al., 2021)."

We have also estimated the sensitivity of our results to the treatment of these phase changes. We found that the conclusions of the article concerning the accuracy and stability of the different SEB schemes hold with a different treatment of phase changes (graphs provided in the response to the review of Michael Lehning). We now address this point in the revised manuscript:

P12 - L329
"Also, as some of the current snowpack and glacier models include the effect of internal phase-change while solving the internal heat equation (e.g. Bartelt and Lehning, 2002, Meyer and Hewitt, 2017), we quantified the sensitivity of our results to this specific treatment of melt/freeze. For that, we have also implemented versions of our three models that include such internal phase-changes in the heat equation."

P16 - L441
"Finally, using the versions of the models including phase-changes in the heat equation, we quantified the sensitivity of these observations to the treatment of the melt/refreeze. While the simulated temperature sometimes differ from our basic implementations (especially in the snowpack test case where melt occurs internally), the general behavior of the models, including the potential presence of instabilities in the Class 2 models, remain unchanged."

136- "SNTHERM (Jordan, 1991), Crocus (Vionnet et al., 2012)"
The typo is corrected.

146- Another step is required if the calculated melt exceeds the available snow mass.
It is indeed important that the local calculated melt does not exceed the available snow mass. In our implementation, if the local melt exceeds the snow mass, layers are locally merged until the melt falls behind the available snow mass. This is now specified in the manuscript.

234- LWout and H are given as examples of fluxes that are nonlinear in the surface temperature; L should also be mentioned as intrinsically nonlinear. H as defined by Eq. (B1) is only nonlinear if C_H is a function of surface temperature. It is, through Ri_b here, but models often neglect this nonlinearity because of the complexity of the resulting derivatives; it is not clear if that is done here. A supplement giving the elements of the Jacobian might be a useful addition.
We added L in the list of SEB terms that are non-linear with respect to the temperature.

In our implementation, we take into account the dependence of C_H to the surface temperature and include its impact on the Jacobian of the system (in order to have a true Newton method with quadratic convergence). Note that not taking this dependence in the Jacobian does not modify the solution of the non-linear system, but only the sequence of iterations and the convergence rate toward this solution.

To make our model readily-available we explicitly wrote the terms of the Jacobian in the new Appendix A.

261- I understand the problem, but I don't understand the benefit of returning the solution to the vicinity of the discontinuity.
The SEB should have a unique solution, but the Newton method is not guaranteed to find it. It can get trapped in a cycle of states around the solution. This situation can be diagnosed from the SEB, but I think that most models just give up and select the last iteration. Does the modified Newton method avoid this problem?
Yes, the goal of the truncation method is precisely to avoid the iterations to be stuck in a loop or to diverge and is quite adapted for the solving of the SEB with a fictitious variable. We've made a Figure as illustration below (the SEB non-linearity has been exaggerated for the illustration). In the case of the standard Newton method without truncation, the break in the slope can send the iterations far from the solution (or into loops depending of

the configuration). In the truncation case, the iteration is moved to the orange point after two truncations. At this point, the Newton scheme can converge normally to the solution.

P9 - L251

"The idea behind truncation is that the Jacobian (i.e. the derivative of the equations with respect to the unknowns to be solved for) computed on one side of a derivative discontinuity does not apply on the other side, and can therefore perturb the convergence towards the solution, typically leading to an endless iteration loop."
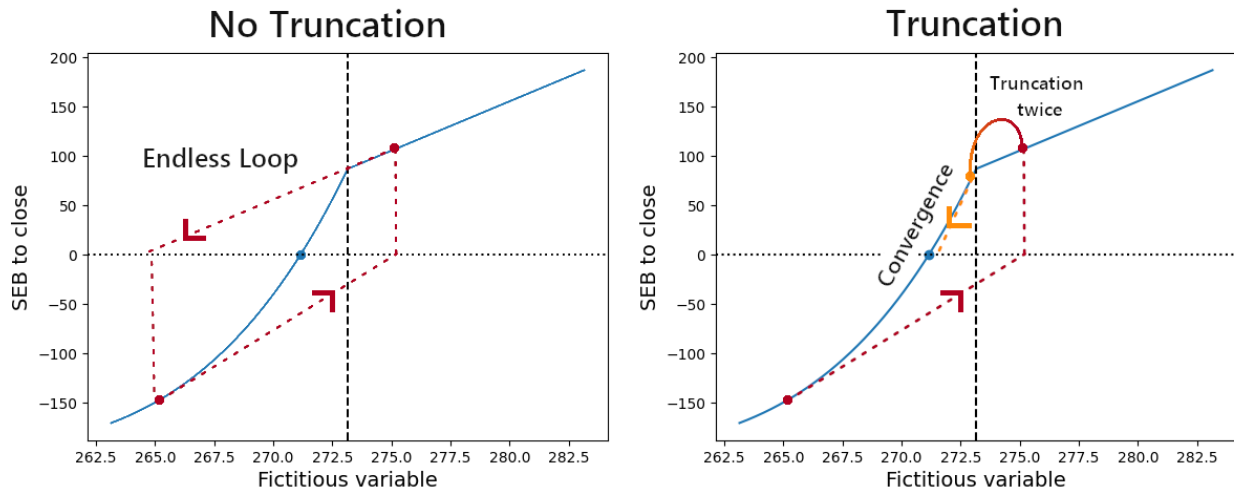

Figure – *Solving of a non-linear SEB with and without a truncation in Newton's method. In the truncation case, the estimation is brought from the red point to the orange point after two successive truncations.*

Note that Newton's method can be made even more robust by applying a truncation at the inflection points. However, this was not done in our case, as the SEB does not displays such inflection point with respect to the surface temperature.

265- Another solution in use, with its own numerical errors, is to linearize the SEB and solve it in one step without iteration (e.g. Best et al. 2011). This is essentially the Penman-Monteith method.
Equation (11) and following
Indeed, some models only solve the linear system with one iteration (for instance Crocus). However in this case, the obtained solution is not the actual backward Euler solution and does not have all its properties. We mention this point in the article.

P9 - L236

"We also note that some models made the choice of performing only a single iteration to solve this linear system of equations (with sometimes an extra iteration to handle specific cases, such as surface melting). However, we chose here to perform multiple iterations, in order to obtain the actual Backward Euler solution."

Be consistent in making diag, up and low superscripts or subscripts.
We corrected the manuscript consistently, with all diag, up, and low being subscripts.

284- "The above equation" is Eq. (13).
We modified the text to state that the "above equation" refers to Eq. (13) and that it allows one to solve the first temperature, as if they were solved with the complete system of Eq. (11).

P11 - L284

"The system of Eqs.(13) is a 2x2 non-linear system where only A_s and B_s need to be re-assembled at each non-linear iteration and whose solution for U_s is the same as the large system of Eqs. (11)."

286- "invert A_diag"

Following a comment of Reviewer 2 n the numerical efficiency of the method, we have proposed to partly rewrite the part of the article detailing the Schur-complement technique. This portion now reads:

P11 - L286

"[…] (ii) compute the products $A_{diag}^{-1} B_{int}$ and $A_{diag}^{-1} A_{up}$ (which is cheaper than directly inverting A_{diag}, (iii) iteratively [...]"

322- "cells which then become"

We wanted to write: "that merges adjacent cells when *they* become smaller than a given threshold". We modified the manuscript accordingly.

379- No refreeze in this test case.

Indeed, in the glacier test case, there is not refreeze as all water is sent to runoff. This will be mentioned in the text.
We still define the phase change rate in terms of melt and refreeze (general definition) and precise that in the glacier test case there is no refreezing.

P14 - L381

"Note that in this specific test case, no refreezing was observed (as melt occurs at the surface and is sent to runoff), meaning that the phase change rate directly corresponds to the melt rate."

421- "Concerning the glacier test-case, Fig. 3 shows"

We corrected the typo.

424- "by about 0.50 K"

We corrected the typo.

440- I have, indeed, seen time step oscillations like this in class 2 simulations. They are not the same as the well-known and catastrophic instability of the explicit Euler method with too large a timestep. Considering the wide use of class 2 models, a stability analysis to understand the origin of these oscillations (not necessarily for this paper) might be of interest.

To better explain the instabilities in Class 2 models we have performed a stability analysis, akin to the ones classically performed for the Forward/Backward Euler scheme. It is provided as the end of this response.
It shows that the standard skin-layer scheme is only conditionally stable, and that there is exist a maximum time step size. The presence of instabilities is favored in the case of large thermal conductivities or of a large derivative of the atmospheric fluxes with respect to the surface temperature in the SEB. On the contrary, these instabilities are hindered in the case of large cell sizes or large specific thermal capacity.
We believe this instability is of the same nature as the one observed with an explicit time-stepping, as it arises from the use of the first internal temperature from the previous time step in the computation of the subsurface heat flux. If the internal temperature from the

current time step is used instead (as in the scheme we propose), this instability is removed.

As mentioned above, the demonstration of the (un)conditional stability of the schemes is now presented in the new Appendix A and discussed in the text.

P21 - L503
"The unstable nature of class 2 models can be shown with a linear stability analysis, provided in Appendix E. Such analysis shows that class 2 models are only conditionally stable, and confirm that instabilities are favored in the case of large time steps and small mesh sizes. We stress that these oscillations can appear even if the time integration of the internal energy budget relies on the Backward Euler method, known for its robustness against instabilities (Fazio, 2001, Butcher, 2008). Our understanding is that the sequential treatment of the standard skin-layer formulation breaks the implicit nature of the time integration by using "lagged" (in other words, explicit rather than implicit) terms. This, combined with the fact that the surface layer does not possess any thermal inertia and that its temperature can thus vary rapidly in time, permits large temperature swings if the time step is too large or the mesh size too small. On the other hand, it can be shown that the two schemes with a tightly-coupled SEB are unconditionally stable (Appendix E), in agreement with the absence of oscillations in their simulations. Notably, the unconditional stability of the coupled-surface scheme proposed in this article entails that the model does not need an adaptive time step size strategy depending on the mesh size. This ensures that it remains robust, regardless of the time step and mesh size.

P26 - L563
"Moreover, a tightly-coupled treatment of the SEB allows unconditional stability, while the standard skin-layer formulation can be unstable and displays large spurious oscillations with large time steps and small mesh sizes."

460- "only marginally worse"
What we wanted to say here, is that sometimes the Class 2 model yield smaller error than the coupled-surface scheme, but when it do so it is only be small margin (which then justifies the use of a coupled-surface model in general). This was visibly not clearly phrased, as Micheal Lehning had the same remark. We rephrased the sentence to:

P17 - L458
"For almost all investigated time steps and in both test cases, the newly proposed scheme displays the lowest level of errors. Sometimes, the class 2 model yields the smallest error, but does so only by a small margin."

We have also re-formulated a similar sentence later in the manuscript.

P20 - L481
"Again, among the three implementations the tightly-coupled surface model yields the smaller errors for almost all investigated mesh refinements (as in the glacier test case, the class 2 model is however sometimes marginally better)."

490- Divergence of the class 2 model from the reference as the mesh is refined in the glacier test case (Fig. 10) is an odd result. I guess that this could happen if the time step in these mesh refinement tests is larger than in the reference. If so, this needs to be stated in the text.

We believe the increase of error with smaller mesh size st a result of numerical instabilities, that develop with small mesh sizes. This is now mentioned in the revised manuscript:

P20 - L490

"Finally, Fig. (10) reveals that in the glacier test case, the phase change rate errors of the class 2 tend to deteriorate with further mesh refinement past a certain point (here for an initial cell number above 90). We interpret this deterioration as a result of the appearance of numerical instabilities that develop with small mesh sizes."

Having said that, it is not apparent that the 225 cell simulation is worse than the one with 45 cells in Fig. 10c.

There are periods in Fig. 10c where the error in the 225 cells simulation is larger than the 45 cells. This is notably the case from mid-June to late-August.

504- "the backward Backward Euler method" sounds like it goes forward. Just one "backward" required.

Indeed. This is now corrected.

509- "mesh size too small"

We corrected the typo.

6.4- Having found G from the SEB, the obvious thing to do in a class 2 model is to use it as a flux boundary condition for the internal temperature calculations. Can any real class 2 model be found that uses the surface temperature as a Dirichlet boundary condition? If not, section 6.4, Fig. 14 and the last sentence of the conclusion should be deleted. A note that this would be the wrong thing to do will suffice.

The potentiality of using the surface temperature as a Dirichlet condition rather than the subsurface conduction flux was made aware to us from reading the publicly available COSIPY code (cosipy/modules/heatEquation.py files, last accessed 08/11/2023) and EBFM codes. However, we stress that these codes use a Forward Euler time stepping, and in this case the using the sub-surface conduction flux or a Dirichlet condition are equivalent.

We think it is important to mention and show that using a Dirichlet condition will lead to greatly deteriorated simulations, as the use of a Dirichlet condition actually numerically stabilizes the system (which can be seen with the absence of instabilities in the orange curve of Fig. 14 and can be demonstrated with a stability analysis, provided at the end of this document). However, this stabilization is at the detriment of accuracy and energy conservation.

We propose to better justify in the manuscript that the use of a Dirichlet condition might be tempting to obtain stability, but that it will produce large errors in response. We have also shorten the first part of the section:

P23 - L511

"As explained in Section 2.2, the heat conduction flux from the surface to the interior of the domain (i.e. G in Equation 3) needs to have the same value in the computation of the SEB and in the computation of the energy budget of the first interior cell. Inconsistencies in G between these two budgets lead to the violation of energy conservation and create an artificial energy source/sink near the surface. Such inconsistencies could be created when implementing the standard skin-layer formulation (class 2 models) due to the sequential

treatment of the surface and internal energy budgets. Indeed, after solving the SEB, one can either use the surface temperature or the subsurface heat flux G as a boundary condition for the computation of the internal temperatures. We note that the use the computed surface temperature as a boundary condition leads to an unconditionally stable numerical scheme (Appendix E). However, using such Dirichlet condition in order to stabilize the standard-skin layer formulation comes at the expense of energy conservation and deteriorates of the simulated results."

Search the text for ", that". In all but one case, it should be "that" or ", which".
This has been corrected.

Albert (1983): https://apps.dtic.mil/sti/pdfs/ADA134893.pdf

Anderson (1968): https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/WR004i001p00019

Best et al. (2011): https://gmd.copernicus.org/articles/4/677/2011/

Clark et al. (2015): https://agupubs.onlinelibrary.wiley.com/doi/full/10.1002/2015WR017200

Dutra et al. (2010): https://doi.org/10.1175/2010JHM1249.1

Yamazaki and Kondo (1990): https://doi.org/10.1175/1520-0450(1990)029<0375:APMFSS>2.0.CO;2