# Responses to RC on manuscript EGUSPHERE-2024-249

***Authors :***
Diego Monteiro, Cécile Caillaud, Matthieu Lafaysse, Adrien Napoly, Mathieu Fructus, Antoinette Alias, Samuel Morin

# 1 Reviewer comments : Emanuel Dutra

**The study presents an improved land surface snow representation in the CNRM-AROME model evaluated over the European Alps at convection-permitting resolution (2.5km). The snow simulations are evaluated against a large sample of in-situ snow depth observations and remote sensing snow cover. The manuscript is well written with good quality graphics and with a very detailed and interesting process-oriented discussion of the results. In my opinion, this manuscript is of interest to the community and fits well with GMD scope. However, there are several decisions on the manuscript organization and results presentation that, in my opinion, limit the main message of the study, and my suggestion would be for the authors to consider some re-organization of the results presentation and discussion, along with a few minor clarifications listed bellow.**

**

We would like to thank Emanuel Dutra for his constructive evaluation of the article. His assessment is positive and detailed, and we appreciate his time in evaluating our work and providing a set of comments for improvement. We have provided a point-by-point response to his comments below.

## 1.1 ES-DIF should not be an experiment but a sensitivity test, It is mentioned in section 2.3.2 : However, note that only one patch is used herein for the NATURE tile, which is not the way the configuration is implemented for the coupled systems CNRM-CM6 and CNRM-ALADIN using 12 patches, and HARMONIE-Climate using 2 patches. and also in the discussion We also note that there are little to no cases where the ES-DIF configuration, with only one patch, has been implemented in an offline or coupled modeling system. I would suggest to re-organize partially the results/discussion having two configurations : D95-3L and ES-DIF-OPT, where ES-DIF-OPT is presented and D95-3L used as benchmark, while ES-DIF, 3-PATCHS, GFLUX, WSN-1 are sensitivity experiment to allow a process-oriented discussion. For example, in my opinion, Figures B1 and B2 are the most interesting results of this study, but are presented in the appendix.

We thank Emanuel Dutra for his suggestions. His remarks tie in with the more general consideration of editorial choices made in the course of writing an article, particularly in the case of research results containing a large number of hypotheses and tested experiments, and the way in which these are to be told and developed.

For this article, we have chosen to focus on three main configurations so as not to overload the reader with a complex set of experiments, allowing us to easily convey the main results we wish to highlight. There are three main results we wanted to focus on :

— The initial configuration doesn't allow to reproduce the seasonal evolution of the snow cover in the European Alps, particularly in areas where vegetation is present. This is due to the lack of an explicit representation of the snowpack and the vegetation (both included in a composite surface layer), and to the way in which the surface energy balance is accounted for in these cases to melt the snowpack.

— A configuration that includes only an enriched representation of the soil and an explicit representation of the snowpack also fails to simulate the seasonal evolution of the snowpack satisfactorily, despite variations in snow height that are closer to observational references than the reference configuration. These limitations are directly linked to the way in which energy balances are calculated when there is a partial snow fraction (this happens much more frequently in the presence of high vegetation, due to the formulation of the snow cover fraction under vegetation).

— A final configuration (including a set of modifications) based on the previous configuration enables to significantly reduce deviations from references by significantly improving the seasonal evolution of snowpack. The modifications tested iteratively allow us to identify the main factors affecting the representation of the snowpack in the ES-DIF.

For this purpose, we have chosen to present the standard configuration used here as a reference, since it is the one used in the previous version of the CNRM-AROME model, and to present the final configuration, since it is the most satisfactory in terms of the metrics evaluated for the representation of snow cover in the European Alps.

We also decided to include the ES-DIF configuration, despite the fact that it is little used in the literature. The choice of this intermediate configuration was based on multiple criteria :

— The distance of this configuration from the others in terms of modifications to the individual elements of the configuration, as well as in terms of snow simulation results.

— The fact that it allows us to highlights difficulties in the representation of the snowpack that are specific to the way energy exchanges at the ground-snow interface are handled (cf. section 4.2). Although the shortcomings of this configuration have already been identified in previous standalone simulation at the point-scale, this study is the first large-scale, coupled surface-atmosphere, spatialized and at high-resolution evaluation of this configuration, and we believe that it offers real added value in terms of its presentation in the body of the text.

— It also underlines the fact that the sudden addition of enriched physical components does not automatically lead to improved metrics, and highlights the importance of the way in which the various components interface and the representation of sub-grid heterogeneities.

For these reasons, ES-DIF appears to us as an intermediate experiment facilitating the understanding of the main messages of the study, and we have chosen in the revised manuscript to keep this experiment in the main body of the text.

The elements communicated above on the main messages of the study also explain why figures B1 and B2 have been placed in appendices. The current organization, which is the result of a choice made with a view to keeping the main body of the text as concise as possible while preserving the main messages of the article, seems to us to be the most optimal. Readers interested in the detailed results of each experiment will find the information in the appendices, and will be able to understand these figures quickly since they are presented in a very similar form in the main body of the text.

Nevertheless, as the reviewer Richard Essery pointed out in his comments too, the last analysis in the appendices C concerning a multivariate comparison at the station appears for the first time in the discussion with almost no introductory sentence to properly apprehend it. To make it easier to understand, we have added it to the body of the text in section 4.4 of the discussion, and provided a brief description to clarify its use in the rest of the paragraph.

### 1.2 The 3-PATCHS : The decision to use 3-patches seems a bit unclear. Why not 2 or 4 or 5 ? Just computational cost ? I understand that this is likely driven also by expert knowledge that is difficult to justify. However, the use of 3-patches, representing bare-soil, low-vegetation, high vegetation has been the approach taken by ECMWF since early 2000s already in ERA-40 (https ://doi.org/10.21957/9aoaspz8) more than 20 year ago.

We would like to thank Emanuel Dutra for his question and remark concerning the choice of the number of patches. As explained in section 2.3.3, this is a trade-off between the representation of the main sub-grid heterogeneities that can be found in three main categories (no vegetation, low vegetation, high vegetation) and the additional computational, writing and storage cost involved in splitting each surface by three. Using a few additional patches would entail additional costs for few benefits (only allowing to separate between high vegetation types), while using only 2 would no longer allow us to distinguish between areas where vegetation is present and those where it is absent. The use of three patches then appeared to us to be the most appropriate and minimal number for the highest improvements.

### 1.3 GLUX : The proposed cap of the conductivity at 5% below a snow fraction of 75% and then linear is a pragmatic approach (mentioned in the discussion), which is an acceptable justification, but more importantly it supports the hypothesis of an overestimation of the ground heat flux, inducing basal melt in partially snow-covered surfaces. The harmonic average between snow/soil conductivities does not account for the air trapped between the snow base and soil top due to living/dead organic matter that will effectively reduce heat conductivity. This is difficult to represent due to the very high spatial variability of such conditions, but I suggest that this is also discussed as a possible explanation for the overestimation of the ground heat flux.

We thank Emanuel Dutra for his suggestion on the potential role of the air trapped by vegetation at the soil-snow interface that can act as a heat conductivity reducer. In all the configurations tested here, this effect is not implemented, and it could be a factor contributing to the overestimation of the heat flux. It could act jointly with other factors not taken into account or poorly modeled, which could lead to a reduction in soil thermal conductivity, such as an underestimation of the organic matter content in the soil surface layers. This is the reason behind the modification of the input databases for soil organic matter content and the activation of its effect, although this had only a negligible effect on the representation of snow cover in our region of interest.

We also note that the effect on thermal conductivity of a layer of organic matter at the soil-snow interface is already available within the SURFEX MEB (Multi Energy Balance) explicit vegetation module (Napoly et al., 2017 ; Boone et al., 2017) and represents a future prospect for coupling with the CNRM-AROME model.

According to this suggestion, we have introduced a sentence at the end of section 4.2 in the revised version of the manuscript to mention the potential effect of factors that would lead the configuration to overestimate thermal conductivity at the soil-snow interface : *"It is important to note that, in addition to the heating feedback from snow-free surfaces to snow-covered surfaces, other factors not directly tested in this study are likely to exacerbate this effect. Indeed, due to their effects on reducing thermal conductivity, some physical processes, such as air trapping at the soil-snow interface due to the presence of a litter layer and/or low vegetation, as well as a poor representation of organic matter content in the upper soil layers, may contribute to and exacerbate basal snowpack melting."*

### 1.4 Impact of initialization : I do not see any reason for the authors not to use the spin-up simulation ? Although it has a reduced impact for the actual period of validation, if the authors have a simulation with land initial conditions that are much better, why not use it ?

We would like to thank Emanuel Dutra for pointing out that we do not use an initial surface state resulting from long spin-up initialization for our experiments. Actually, we would also have preferred to run the simulations with a correctly initialized surface, but technical and computational cost considerations constrained our choice.

In fact, the spin-up was initially carried out over few months for the D95-3L configuration, with little impact on the initial soil conditions, since a "force-restore" scheme was in use, thus requiring no particular spin-up. The implementation of the new ES-DIF configuration using a multi-layer soil with explicit diffusion implies a spin-up of at least 5 to 10 years to reach a point of equilibrium in terms of integrated water and heat content of the soil columns (Christensen, 1999 ; Cosgrove et al., 2003). The main problem is that the numerical cost of the simulations (very high in the case of simulation with the CNRM-AROME model on the ALP-3 domain) limited the realization of a coupled surface-atmosphere spin-up over several years.

In order to avoid the high numerical cost of a coupled spin-up, we decided to perform a standalone simulation of the surface, forced by atmospheric fields from previous CNRM-AROME simulations carried out as part of CORDEX's Convection FPS (Coppola et al., 2020 ; Pichelli et al., 2021 ; Caillaud et al., 2021). However, these standalone simulations were only set up after most of the experiments presented in the study had been carried out. So as not to have to run all the experiments again, and to be able to maintain equivalent initial conditions for all the experiments, we decided to continue with the default initialization, while evaluating the impact on snow cover of a surface with a long spin-up. The differences both initializations being negligible, as shown in section 2.5, we believe that, although not optimal, the lack of a long spin-up does not affect the results and conclusions of our study. We also considered that the added-value of running the simulations again was not worth the resources required to run the experiments on the High-Performance Computing (HPC) infrastructure.

### 1.5 Main result not much explored/discussed : The sensitivity results in Figure B1, in particular the activation of the 3-patches shows that even at convection-permitting resolution of 2.5km representing the sub-grid scale variability of land surface heterogeneities of land cover are fundamental. I think that this is a very interesting result in particular within the current European destination earth program, showing that increasing horizontal resolution is not enough for a better representation of the land surface processes mostly due to the very high spatial scale of surface heterogeneities that impact the mean state over the grid-box. I would suggest the author discuss a bit on this in the conclusions and even mention it in the abstract, if they agree.

We thank Emanuel Dutra for his remark and suggestion. We agree that this is an interesting results of the study. Even at kilometric resolutions, at least in the case of CNRM-AROME, doing without a representation of sub-grid heterogeneities of the surface implies numerous errors, particularly concerning the estimation of the grid-average energy balance, with a significant effect on the simulation of snow cover in mountain areas.

It should be noted, however, that while we believe these conclusions to be extendable to other surface variables and other regions, the effect of representing sub-grid heterogeneities is in the specific case of snow simulation exacerbated by poor heat flux management due to the assumption of a shared soil between snow-covered and snow-free areas. In the absence of this assumption, with, for example, two separate soil columns for snow-covered and snow-free surface, we can expect a considerable reduction in the overestimation of basal heat flux and a clear improvement of the snow cover simulation, even using only one patch.

Based on the suggestion of Emanuel Dutra, we highlighted this result in the conclusion L.780-785 in the revised manuscript : *"Their effectiveness confirms the hypotheses put forward to explain the exaggerated melting in the ES-DIF simulation and underlines the importance, even at kilometer resolution, of taking into account the main sub-grid heterogeneities concerning surface type in mountainous terrain. In the analysis of the preferred configuration ES-DIF-OPT, this allows the simulation of snow cover to be satisfactory compared with the references used in many cases."*

And a sentence was modified in the abstract L.15-17 in the revised manuscript : *"These limitations are addressed in further configurations that highlight the importance, even at kilometer resolution, of taking into account the main sub-grid surface heterogeneities and improving representations of interactions between fractional snow cover and vegetation."*

### 1.6 Line 84 : e.g. HTESSEL (Balsamo et. al. 2009) : Suggest to change to ECLand (https ://www.mdpi.com/2073-4433/12/6/723) as this is an updated reference for the ECMWF model that also has a multilayer snow scheme (https ://doi.org/10.1029/2019MS001725)

The manuscript has been revised accordingly.

### 1.7 Lines 87-88 : The sentence The identify shortcomings may also explained some of the snow cover issues raised in  is very unclear, please rephrase it.

The sentence have been rephrased in the revised manuscript : *"Moreover, the factors proposed in the study to explain the erroneous representation of the snowpack in CNRM-AROME are strongly suspected to contribute to the shortcomings in the representation of seasonal snow cover documented in coarser resolution coupled simulations using SURFEX-ISBA LSM, such as CNRM-ALADIN (Termonia et al., 2018) in the Alps (Monteiro and Morin, 2023) and CNRM-CM6 in high-latitude boreal forests (Decharme et al., 2019)."*

### 1.8 Eq. (5) : Please provide in text (or table) and typical z0 values in the presence of high vegetation used in the model to help understand actual behavior of the parameterization of snow cover fraction in these situations. THis also links with the WSN factor change from 5 to 1. Using a typical z0 value, showing a figure of the snow cover fraction as a function of snow depth with WSN=5 and WSN=1 would be illustrative.

We thank Emanuel Dutra for his suggestion on additional figures and tables that will help interested readers to understand the behaviour of the snow cover fraction parameterization over vegetation.

The roughness length involved in the formulation is that of the average roughness length over the nature tile for a given grid cell, meaning that it includes not only the vegetation but results from an aggregation of all surface types inside the grid cell, snow included. Therefore its value varies along the winter season, with the amount of snow on the ground, and monthly with the LAI of the vegetation.

Figure 1, added in appendix A in the revised manuscript is a map of the mean surface roughness length values over the two winter seasons (November to April for the 2018-2019 and 2019-2020 periods) using the D95-3L configuration.

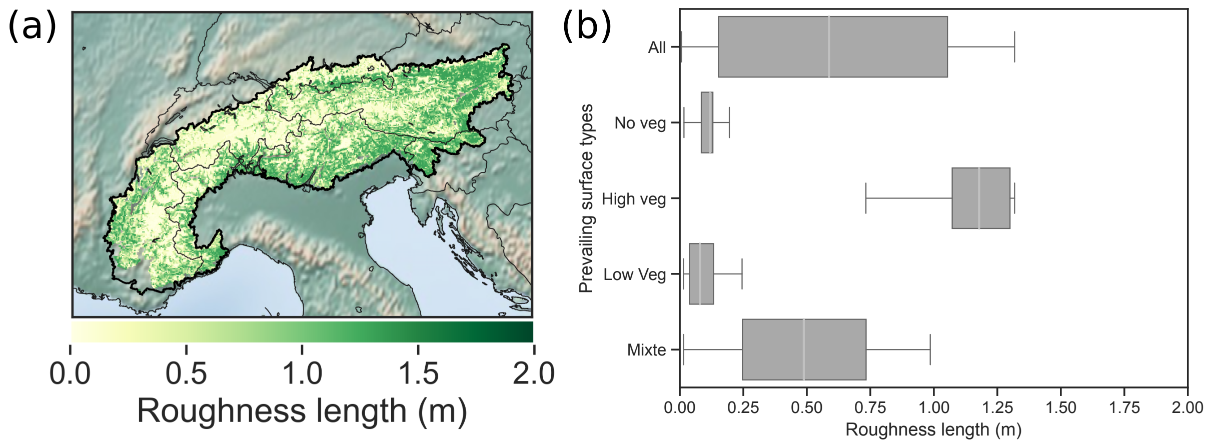


FIGURE 1 – **(a)** Map of the mean surface roughness length values over the two winter seasons (November to April for the 2018-2019 and 2019-2020 periods) using the D95-3L configuration. **(b)** Boxplot representing the spatial distribution of the mean roughness length values over the two winter seasons (November to April for the 2018-2019 and 2019-2020 periods) using the D95-3L configuration classified by prevailing type of surface (see section **??** for details). It is noteworthy that, although the roughness length values are displayed for configuration D95-3L and vary with the amount of snow on the grid cell, they are very similar for all the configurations tested.

Figure 2, added in appendix A in the revised manuscript show the snow cover fraction as a function of snow depth for two combinations of different roughness length $z0$ at a given scaling factor $Wsn$, and two combination of the scaling factor $Wsn$ for a given roughness length $z0$.

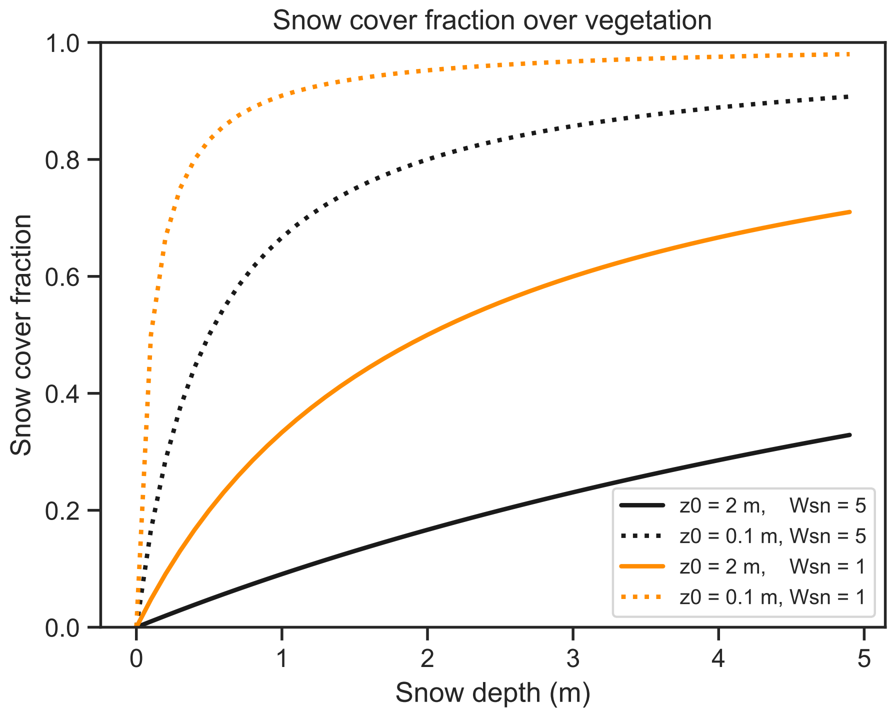


FIGURE 2 – Snow cover fraction over vegetation as a function of snow depth for multiple combination of values for the roughness length $z0$ and the scaling factor $Wsn$. The black curves represent the sensitivity of the snow cover fraction parameterization using a value of $Wsn = 5$ as is the case for the D95-3L and ES-DIF configuration, while the orange curves a value of $Wsn = 1$, used for the ES-DIF-OPT configuration.

### 1.9 Table 1 : Table presents in the last line computational time relative to D95-3L I must admit that I was very surprised to see such an increase when activating ES-DIF (+15%), and even more with -OPT just with 3 patches. In NWP/climate models the land surface component is normally a rather negligible part in the computational cost due to the simplicity of the calculations and to the 1D nature that typically fits very well MPI /OpenMP implementations. This is mostly a curiosity, but do the authors have some explanation for such a significant computational cost increase ?

We thank Emanuel Dutra for his question concerning the increased computational cost of the successive changes of the land surface configuration.

Indeed, the increased computational cost of the surface is small compared to the absolute computational cost of the atmospheric part. It turns out that, if we look more closely, the dominant factor accounting for most of the increase lies in the time steps used to write the model outputs (every hour). Although the computational overhead is not high, the use of multi-layer schemes for snow and soil, followed by the activation of three patches, multiplies the number of prognostic and diagnostic variables written each hour. Each hour, when switching to the ES-DIF configuration, the number of prognostic variables for soil alone is multiplied by 7 (from 2 to 14 layers), while that for snow is multiplied by 12 (from 1 to 12 layers). In addition, all prognostic and diagnostic variables are multiplied by three when 3 patches are activated. This highlights that increasing the realism of the land surfaces, hence the number of variables, could be limited for some applications by input/output considerations such as the writing of prognostic and diagnostic files.

However, the values shown in Table 2 of the original manuscript may not reflect the additional computational cost that will be of interest to readers, given that input/output processing can be considerably optimized and is workflow-dependent. In the revised manuscript, the computational costs have been replaced to take into account only the model execution time step (excluding input/output considerations).

### 1.10 Line 424 ; sigma standard deviation is not used in the equations, no need to define it

The manuscript has been revised accordingly.

## Références

Boone A, Samuelsson P, Gollvik S, Napoly A, Jarlan L, Brun E, Decharme B (2017) The interactions between soil–biosphere–atmosphere land surface model with a multi-energy balance (isba-meb) option in surfexv8–part 1 : Model description. Geoscientific Model Development 10(2) :843–872

Caillaud C, Somot S, Alias A, Bernard-Bouissières I, Fumière Q, Laurantin O, Seity Y, Ducrocq V (2021) Modelling mediterranean heavy precipitation events at climate scale : an object-oriented evaluation of the cnrm-arome convection-permitting regional climate model. Climate Dynamics 56(5) :1717–1752, DOI 10.1007/s00382-020-05558-y

Christensen OB (1999) Relaxation of soil variables in a regional climate model. Tellus A 51(5) :674–685

Coppola E, Sobolowski S, Pichelli E, Raffaele F, Ahrens B, Anders I, Ban N, Bastin S, Belda M, Belusic D, et al. (2020) A first-of-its-kind multi-model convection permitting ensemble for investigating convective phenomena over Europe and the Mediterranean. Climate Dynamics 55(1) :3–34

Cosgrove BA, Lohmann D, Mitchell KE, Houser PR, Wood EF, Schaake JC, Robock A, Sheffield J, Duan Q, Luo L, et al. (2003) Land surface model spin-up behavior in the north american land data assimilation system (nldas). Journal of Geophysical Research : Atmospheres 108(D22)

Decharme B, Delire C, Minvielle M, Colin J, Vergnes JP, Alias A, Saint-Martin D, Séférian R, Sénési S, Voldoire A (2019) Recent changes in the ISBA-CTRIP land surface system for use in the CNRM-CM6 climate model and in global off-line hydrological applications. Journal of Advances in Modeling Earth Systems 11(5) :1207–1252

Monteiro D, Morin S (2023) Multi-decadal analysis of past winter temperature, precipitation and snow cover data in the european alps from reanalyses, climate models and observational datasets. The Cryosphere 17(8) :3617–3660, DOI 10.5194/tc-17-3617-2023, URL `https://tc.copernicus.org/articles/17/3617/2023/`

Napoly A, Boone A, Samuelsson P, Gollvik S, Martin E, Seferian R, Carrer D, Decharme B, Jarlan L (2017) The interactions between soil–biosphere–atmosphere (isba) land surface model multi-energy balance (meb) option in surfexv8–part 2 : Introduction of a litter formulation and model evaluation for local-scale forest sites. Geoscientific Model Development 10(4) :1621–1644

Pichelli E, Coppola E, Sobolowski S, Ban N, Giorgi F, Stocchi P, Alias A, Belušić D, Berthou S, Caillaud C, et al. (2021) The first multi-model ensemble of regional climate simulations at kilometer-scale resolution part 2 : historical and future simulations of precipitation. Climate Dynamics 56(11) :3581–3602, DOI 10.1007/s00382-021-05657-4

Termonia P, Fischer C, Bazile E, Bouyssel F, Brožková R, Bénard P, Bochenek B, Degrauwe D, Derková M, El Khatib R, Hamdi R, Mašek J, Pottier P, Pristov N, Seity Y, Smolíková P, Španiel O, Tudor M, Wang Y, Wittmann C, Joly A (2018) The ALADIN System and its canonical model configurations AROME CY41T1 and ALARO CY40T1. Geoscientific Model Development 11(1) :257–281, DOI 10.5194/gmd-11-257-2018