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 1 : Richard Essery

In addition to being able to develop and advect convective precipitation with the model dynamics, better resolution of topography and snow cover have been proposed as benefits of modelling at convection permitting scales. Monteiro et al. investigate the benefits of improvements in the land surface and snow components of a convection-permitting regional climate model. The D95 configuration used as the baseline is far from the state of the art in climate and NWP models, but it is relevant as the land surface model in the current CNRM-AROME regional climate model.

**

We thank Richard Essery for the insightful comments and suggestions, which have lead to improving the quality of our manuscript. Below we provide point-by-point answers to the suggestions and comments.

1.1 Abstract

1.1.1 From the abstract alone, it is not clear what multiple patches for land surface grid points means.

We thank the reviewer for this comment. Indeed, it may not be explicit what multiple patches refers to for readers beyond the SURFEX-ISBA community. Accordingly, we modified the sentence l.11-13 : *"More specifically, the study tests the influence of various changes in the land surface configuration, such as the use of a multilayer soil and snow scheme, the division of the energy balance calculation by surface type within a grid cell (multiple patches), new physiographic databases and parameter adjustments."*

1.2 Introduction

We thank Richard Essery for the two following comments concerning the developments of the introduction, for which some formulations can be misleading in their current form.

1.2.1 the use of high-resolution models would minimize modeling uncertainties, by limiting the use of sub-grid parameterizations ; with the exception of convection, none of the subgrid phenomena listed earlier in this paragraph are parametrized in AROME.

This sentence did not refer directly to the atmospheric processes listed above, but was intended to emphasize that higher resolution is also of interest for the representation of surface heterogeneities, limiting the need for parameterizations to take into account sub-grid heterogeneities in LSM (Land Surface Model). Nevertheless, the succession of paragraphs can lead to confusion, and consequently sentence L.38-39 has been deleted, and a sentence L.47 has been added : *"In addition to their capacity to explicitly resolve deep convection and thereby enhance the representation of precipitation extremes (Caillaud et al., 2021), models operating at the km-scale make it possible to better represent the topography of mountain areas, and the heterogeneities that characterize the surface through higher resolution, holding great potential for mountain regions."*

1.2.2 Assessing the representation of snow in coupled configurations is a necessary complementary approach to standalone model run. I would rather put it the other way round : coupled configurations are essential for NWP and climate applications, and it is standalone model runs that can be a useful complement to the evaluation. Raleigh et al. (2015) and Lapo et al. (2015) cited here do not consider coupled surface-atmosphere simulations.

We agree that the proper way to express the idea would be the other way around. As the sentence can be misleading we decided to rephrased it in the revised manuscript L.31-32 : "Testing snow models in standalone "offline" configurations is not sufficient, and tests in coupled configurations are required. This is particularly challenging in mountainous areas."

With regard to the use of the citations Raleigh et al. (2015) and Lapo et al. (2015) the objective was to provide evidence that, under coupled conditions, atmospheric errors (an unavoidable consequence of using a model) can significantly affect the accuracy of snowpack simulations, even in the absence of coupling. This is precisely what the two articles demonstrate by measuring the impact of errors in atmospheric forcings on snow simulations.

1.3 2.2.1

1.3.1 A display equation should be integrated into the text as part of a sentence ; equations (1) (6) are not. latent L heat flux should be LE to match equation (4.2). LWu in equation (3) really is incorrect ; it should include reflected LWd if the surface emissivity is not 1 (Kirchoffs law of thermal radiation). the atmospheric fluxes received for all tiles and patches means specifically the incoming radiation and precipitation fluxes.

We would like to thank Richard Essery for pointing out the formulation errors in the display equations, which have been incorporated into the sentences in the revised version.

We have also corrected the formulation error in equation (3) to take account of the LWd reflected :

$$Rn = SWd + \varepsilon_s LWd - SWu - LWu \tag{1}$$

with Rn the radiation balance ($\mathrm{W\,m^{-2}}$), SWd the incoming solar radiation, LWd the infrared atmospherical radiation, SWu the reflected shortwave radiation and LWu the emitted longwave radiation.

1.4 2.3

1.4.1 The French expérience should be translated as experiment The density, an exponentially decreasing function, forced to 100 kg m3 for fresh snow, limited to 300 kg m3 for aged snow ; density is an increasing function. 300 kg m3 can be low for aged snow and could contribute to overestimates of depth.

We thank Richard Essery for pointing out the typo, which we have corrected accordingly.

We would also like to thank him for his insightful comment regarding the role of density formulation in the overestimation of simulated snow height. While we have not tested this element, it is a potential factor that may contribute to the modeled overestimation of snow height in the D95-3L configuration. However, our analysis suggests that the primary factor is likely the underestimation of melt, due to his formulation driven by the daily variations of the surface temperature in the presence of vegetation, as discussed in section 4.1.

1.5 2.4

1.5.1 The horizontal resolution of ERA5 is 0.25°, not 50 km ; was it regridded ?

The ERA5 data have indeed been regridded, and the method employed is quadratic interpolation. The reason for the use of these forcings instead of the native ones is purely technical : they are the highest resolution forcings initially available on the servers at the time the simulations were initiated. Given the time constraints associated with downloading data at native resolution, it was preferable to use those available.

However, the impact of using these instead of finer resolution may be significant, particularly on the formation of fine-scale structures close to the domain boundaries. Indeed, the literature on the subject recommends that the resolution jump for dynamic downscaling should not exceed a factor of 12 (Leduc and Laprise, 2009 ; Leduc et al., 2011). In practice, the greater the jump, the greater the distance from the edges of the domain required for the correct formation of fine-scale structures (i.e., spatial spin-up) to be respected, and therefore the greater the number of points at the edges that will be unusable (Matte et al., 2017).

In the case of our simulation on the ALP-3 domain, the European Alps are located at the center of the domain, several hundred kilometers (more than a hundred grid points) from the domain boundaries, we therefore consider that spatial spin-up is effective and does not affect the results of our simulations.

1.6 Figure 3

The distribution of elevation in the Alpine domain could be added to the bar chart as an indicator of representativeness.

Following the recommendation of Richard Essery, the frequency distribution of elevation, as represented by a digital elevation model (DEM) at a resolution of 100 m, has been incorporated into the bar chart in

Figure 3 of the revised manuscript. Although we believe that these data can be useful to readers, as they allow for the comparison of the elevational distribution of available observations to the areal elevational distribution of our region of interest, they do not provide direct insights into the representativeness of observation stations. It is important to note that while the elevation distribution of observations may be similar to the frequency distribution of the actual elevations, this does not necessarily indicate an accurate representation of their spatial distributions. This could potentially lead to an over- or under-sampling of certain areas, but it does not provide insight into the reliability of statistics or scores calculated over a range of elevations.

1.7 2.6.2

1.7.1 State that the MODIS product used is MOD10A1F. How will gap filling and dense vegetation influence uncertainties in snow cover duration ? To determine snow cover duration, a threshold is put on observed snow cover fraction and modelled snow depth, but the model already calculates a snow cover fraction (equations 4 6) ; why not use that ? Are the observation and model thresholds consistent ?

According to (Hall et al., 2019), the gapfilling algorithm replaces the value of pixels categorized as cloud for a given date with the NDSI value of the same pixel for the past nearest date for which it is available. According to the article, values are found in most cases within the previous 3 days (on a study area in the US Rockies), up to 10 days at most. Uncertainties therefore need to be considered on a daily scale regarding NDSI variation for a given pixel (in the case of new snowfall leading to an increase in NDSI, or melting of the snowpack leading to a decrease), which may therefore not be seen.

Working on an aggregated indicator such as snow duration, based on a binary snow/no snow classification determined using a relatively low NDSI value (i.e. 0.2), we expect the impact of daily NDSI uncertainties on snow duration estimation to be small in most cases.

Indeed, if we consider the uncertainties at the beginning and/or end of the season regarding the duration of snow cover, we can expect them to be of the order of 3 to 10 days at most (i.e. order of magnitude of the uncertainties on the daily NDSI). They can therefore be high in the worst cases (from 10 to 30%) for elevation below 1500 m, where the duration of snow cover is of the order of 25-30 days (cf. figure 1), but do not exceed 10% in the most unfavorable cases at 1500 m and above (median duration of snow cover of the order of 115 days, cf. figure 1).

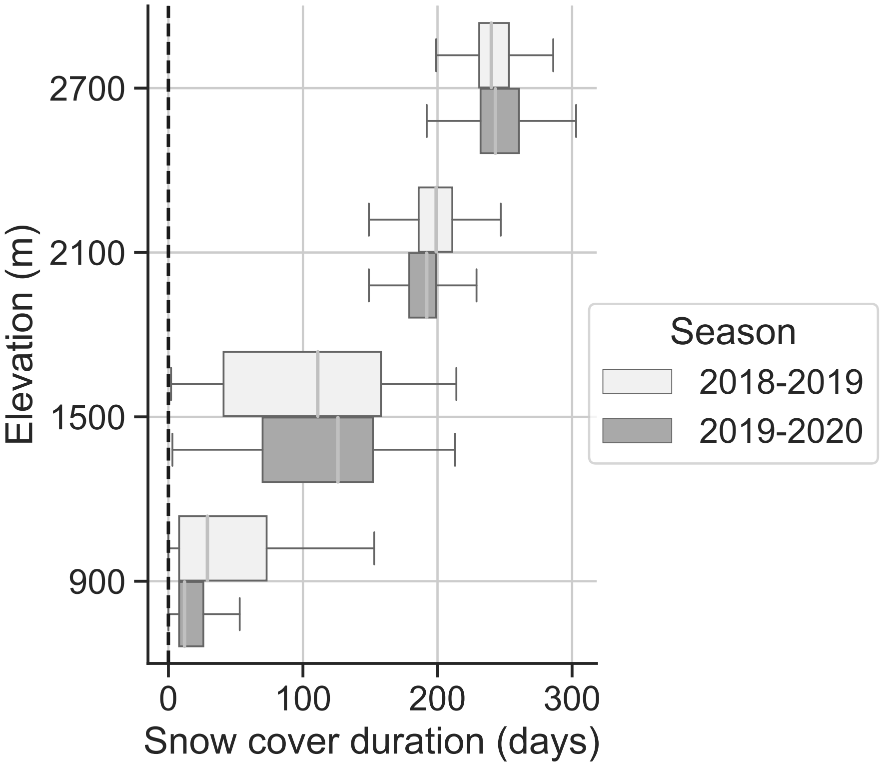

Figure 1 – Boxplot representing the spatial distribution of the snow cover duration values for two seasons (2018-2019 and 2019-2020) for multiple elevation bands in the European Alps.

Regarding the uncertainties in forest areas, it's true that, since they are based on optical detection, the masking effect of the canopy attenuates the signal and thus the NDSI as compared to an equivalent quantity of snow in an open area. Several studies have evaluated MODIS products from the NSIDC Collection 5, directly presented as a binary (value (0/1)) indicating for each pixel the presence or absence of snow, constructed from an NDSI threshold, variable between 0.1 and 0.4 as a decreasing function of NDVI (spectral index depending on the presence of vegetation). On these products, Parajka et al. (2012) was able to show by comparison with in situ snow depth measurements in a mountainous watershed in Slovakia that they maintain an accuracy score of over 81% for mixed forest and open areas, and close to 92% for forest-only areas, compared with 98% in open areas. Gascoin et al. (2015) in a Pyrenean watershed show accuracies higher than 90% in forested areas when compared with in situ snow height measurements and satellite observations at higher resolutions (i.e. Landsat 5 and 7). We therefore have good confidence in the ability of MODIS sensors to detect snow under forest cover, provided that we lower the NDSI threshold.

In our study, we used Collection 6 of the NSIDC MODIS product, in which only NDSI values are provided. A snow presence/absence binary variable was constructed, using an NDSI threshold of 0.2 for all areas, corresponding to a relatively low snow fraction (of the order of 20-30% according to Salomonson and Appel (2004)) to facilitate the detection of snow beneath the forest. In a previous article (Monteiro and Morin, 2023), we were able to evaluate the product by comparing the absence and presence of snow in MODIS pixels with in situ snow depth measurements, taking different snow depth thresholds for a given NDSI threshold. Several scores on the duration snow cover were calculated, as well as the confusion matrix based on the absence and presence of snow. The accuracy scores showed values above 90% for snow depth thresholds ranging from 1 to 5 cm, and the differences in snow cover duration show that the 1 cm threshold minimizes mean absolute errors.

It should also be noted that we use these snow cover durations as comparative values (order of magnitude) and not as absolute references, which we should approach as closely as possible. As the differences between the simulated snow cover durations and those calculated from satellite observations are very large compared with the uncertainties of the snow cover durations estimated by satellite observations, we consider that these errors have little impact on the main conclusions of our study.

Regarding the last part of the question, we have decided not to use the snow fraction as calculated by the model, as we consider that the function which determines it is much more an adjustment variable used to modulate fluxes at the soil-atmosphere interface (which should be modified in the future), than a realistic estimate of the snow cover fraction within a pixel.

1.8 2.7.4

1.8.1 σ_x et σ_y are not in the equations. x and y are a model variable and an observation in some order. r_{xy} here becomes R in 3.1.

We thank Richard Essery for his remarks, the manuscript have been corrected accordingly.

1.9 3.1

1.9.1 Figure 6 suggest that ME and R^2 are poor indicators of model performance. A low value of ME can be obtained by averaging underestimates and overestimates at different times. The D95-3L simulations are very poor for the melt season, but are given R^2 values exceeding 0.8 in all cases.

We agree with Richard Essery that these scores exhibit a number of limitations.

In the case of the mean error, it is possible to obtain a score close to zero, despite the presence of over- and underestimation of a similar magnitude. This phenomenon occurs when the two types of estimation offset each other, resulting in a net zero error. Nevertheless, the mean error remains an relevant score that can highlight systematic biases (over- and/or under-estimation) by providing a straightforward and easily comprehensible indication of the sign of differences.

In the specific case of this study, where the differences are systematic in the vast majority of cases considered, the score can be used to quantify (in conjunction with the simulated and observed snow depth curves) the amplitude of overestimations. Concerning the R^2, in our opinion it illustrates well, in the same way as the mean error, the differences we wish to highlight between the curves. As Richard Essery mentionned, the R^2 remains high (always above 0.8) even in the case of the D95-3L configuration, where a clear discrepancy at the end of the season can be observed. Nevertheless, its high score serves as a reminder that, until this late-season discrepancy, the snow model's behavior regarding the accumulation and evolution during the winter months remains satisfactory, and for all the configurations tested.

We have therefore chosen to retain these two indicators in the revised version of the manuscript.

1.10 3.2

1.10.1 Observed SCD per elevation band would be a useful complement to Figure 7. Errors can be larger at intermediate elevation simply because there is more room for error when SCD is not close to 0 or 365 days.

We thank Richard Essery for his relevant remark pointing out a potential factor to explain the largest differences of the snow cover duration that we obtain at intermediate elevations.

As shown in Figure 1, over the four elevation bands investigated, the variance in snow cover duration is highest at intermediate elevations (i.e. 1500 m). However, it is significantly reduced at 2100 m and 2700 m for both seasons. Nevertheless, the errors calculated for the different configurations are in many cases similar or even greater for the 2100 m elevation band than for the 1500 m elevation band.

Although we believe this may be a factor that mechanically widens errors at intermediate altitudes, it probably acts to second order compared to the presence of a partial snow cover fraction exacerbated by the presence of vegetation as discussed in section 4.2 of the manuscript.

Based on these remarks, we have decided to add the distribution of snow cover duration values for each of the two years in Figure 4 of the revised manuscript.

1.11 Discussion

1.11.1 The amount of the Discussion dedicated to discussing results in appendices that have not yet been presented to the reader is odd. If the results are important enough to discuss in detail, they should be presented in sequence in the main text.

We would like to thank Richard Essery for his valuable insights on the content of the appendices and the discussion. Consequently, figure C2 in appendix C has been relocated to the discussion section, accompanied by a more comprehensive description of the figure in the revised manuscript. We believe these amendments will have a beneficial impact on section 4.4, enhancing the clarity and accessibility for readers.

1.12 4.2

1.12.1 ES-DIF R^2 is only degraded compared to D95-3L at 2100 m in Figure 6. Evaluating the ES-DIF configuration that is not actually used in offline or coupled simulations is of limited value ; it has already been effectively rejected.

We would like to thank Richard Essery for his comments on the discussion of the ES-DIF configuration results.

With regard to the first remark concerning the degradation of the R^2 score between the ES-DIF and D95-3L configurations, which is only present at 2100 m and not at all elevations, this was a typo error that has been corrected in the revised manuscript.

However, we do not share the referee's view expressed in the second comment that the ES-DIF configuration is not used in offline and coupled configurations and that its evaluation would therefore be of limited value. It is true that, as presented, the ES-DIF configuration is not currently used in NWP nor climate contexts. However, the ES multi-layer snow model and the DIF soil model are used in many contexts, from global climate simulation with CNRM-CM6 (Decharme et al., 2019) to high-resolution regional simulations with the HARMONIE-AROME model (Belušić et al., 2020). It's also true that this configuration has already been evaluated with weaknesses pointed out by (Napoly et al., 2020 ; Nousu et al., 2023), but in a stand-alone point-scale evaluation. To the best of our knowledge, our study is the only one to evaluate such configuration in a coupled, spatialized at high-resolution context in a mountain region. In addition, autonomous evaluation of the ES-DIF configuration (i.e. without the adjustments made in ES-DIF-OPT) makes it possible to document its weaknesses and iteratively test the various modifications needed to mitigate them.

Références

Belušić D, de Vries H, Dobler A, Landgren O, Lind P, Lindstedt D, Pedersen RA, Sánchez-Perrino JC, Toivonen E, van Ulft B, Wang F, Andrae U, Batrak Y, Kjellström E, Lenderink G, Nikulin G, Pietikäinen JP, Rodríguez-Camino E, Samuelsson P, van Meijgaard E, Wu M (2020) Hclim38 : a flexible regional climate model applicable for different climate zones from coarse to convection-permitting scales. Geoscientific Model Development 13(3) :1311–1333, DOI 10.5194/gmd-13-1311-2020, URL `https://gmd.copernicus.org/articles/13/1311/2020/`

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

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

Gascoin S, Hagolle O, Huc M, Jarlan L, Dejoux JF, Szczypta C, Marti R, Sánchez R (2015) A snow cover climatology for the pyrenees from modis snow products. Hydrology and Earth System Sciences 19(5) :2337–2351, DOI 10.5194/hess-19-2337-2015

Hall DK, Riggs GA, DiGirolamo NE, Román MO (2019) Evaluation of modis and viirs cloud-gap-filled snow-cover products for production of an earth science data record. Hydrology and Earth System Sciences 23(12) :5227–5241

Lapo KE, Hinkelman LM, Raleigh MS, Lundquist JD (2015) Impact of errors in the downwelling irradiances on simulations of snow water equivalent, snow surface temperature, and the snow energy balance. Water resources research 51(3) :1649–1670, DOI 10.1002/2014WR016259

Leduc M, Laprise R (2009) Regional climate model sensitivity to domain size. Climate Dynamics 32 :833–854

Leduc M, Laprise R, Moretti-Poisson M, Morin JP (2011) Sensitivity to domain size of mid-latitude summer simulations with a regional climate model. Climate dynamics 37 :343–356

Matte D, Laprise R, Thériault JM, Lucas-Picher P (2017) Spatial spin-up of fine scales in a regional climate model simulation driven by low-resolution boundary conditions. Climate Dynamics 49 :563–574

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, Welfringer T (2020) Isba-meb (surfex v8. 1) : model snow evaluation for local-scale forest sites. Geoscientific Model Development 13(12) :6523–6545

Nousu JP, Lafaysse M, Mazzotti G, Ala-aho P, Marttila H, Cluzet B, Aurela M, Lohila A, Kolari P, Boone A, et al. (2023) Modelling snowpack dynamics and surface energy budget in boreal and subarctic peatlands and forests. EGUsphere 2023 :1–52

Parajka J, Holko L, Kostka Z, Blöschl G (2012) Modis snow cover mapping accuracy in a small mountain catchment–comparison between open and forest sites. Hydrology and Earth System Sciences 16(7) :2365–2377

Raleigh M, Lundquist J, Clark M (2015) Exploring the impact of forcing error characteristics on physically based snow simulations within a global sensitivity analysis framework. Hydrology and Earth System Sciences 19(7) :3153–3179

Salomonson VV, Appel I (2004) Estimating fractional snow cover from modis using the normalized difference snow index. Remote sensing of environment 89(3) :351–360, DOI 10.1016/j.rse.2003.10.016