REVIEW 1

Review of "Snow cover prediction in the Italian Central Apennines using weather forecast and snowpack numerical models", by Raparelli et al.

Summary

This paper presents snow cover simulations at 3 km resolution over the Apennines mountain range in Italy, during one winter season, using two very different land surface models: the Noah model including a simple one-layer snowpack scheme resolving the mass and energy budget only, and the Alpine3D model including the complex SNOWPACK model resolving the microstructure of the snowpack. Noah is coupled to the WRF meteorological forecast model, which is also used to drive Alpine3D simulations. The authors assess these simulations in terms of near-surface meteorological variables, snow height and daily snow height variations (against in situ measurements) and in terms of snow cover extent (against MODIS-derived snow cover fraction products). They show lower snow height biases for WRF-Alpine3D but a slightly better representation of snow cover area when compared to satellite products.
Many papers in the past years have analyzed kilometer resolution snow cover simulations in mountains, driven by weather forecasts or analyses. This paper is interesting as it focuses on a lower latitude mountain range, in a different climate and lower elevations. It also compares to different models. However, in its present form, it doesn't bring much novelty compared to past studies and could push much further the comparison between simple and complex snow cover models. Therefore, I would recommend major revisions before publication.

We thank the reviewer for the time spent evaluating our work and we thank him/her for the useful and constructive comments. We have answered all comments. Note that our answers are in **BLUE** in the following text.

Major comments

- This paper follows the methodology of several previous papers, and in particular Quéno et al. (2016). To bring more novelty compared to previous papers, it should focus more on the particularity of this study and develop a deeper analysis of its originalities (mostly snow cover simulations in milder climate, and comparison of different complexity models).
As the reviewer suggests, these themes are a peculiarity of this paper. Thus, they will be further investigated and developed in the next version of the manuscript.

- The assessment of the model simulations over only one winter (3 months actually) can be problematic. There is a strong interannual variability between winter seasons, especially in terms of snow cover. Adding some more winter seasons to the comparison would increase the significance of the evaluation (e.g., Essery et al., 2013).
We will extend model simulations to winters 2019-2020 and 2020-2021, thus covering three entire winter seasons. Winter 2019-2020 was dry and mild, while winter 2020-2021 was much snowier than the previous two, thus the models are evaluated over winters with large interannual variability.

- A more in-depth analysis of the processes driving the snowpack evolution is necessary to highlight differences between simple and complex snowpack models. Particularly, many hypotheses about the better representation of processes by WRF-Alpine3D compared to WRF-Noah are not proved. Pushing further that analysis would add significance to the paper. It would also provide elements for a more advanced discussion of simple snow height or snow cover extent statistics.

To better understand the differences between the models in reproducing the snowpack evolution, the authors also considered snow density data, obtained from other sources. Knowing snow density, we were able to compare measured and simulated snow water equivalent data, and thus to assess the contribution of compaction and melting to snow settlement and the ability of the models to reproduce the observations.

Specific comments

- Abstract: the results should be more summarized, with less numbers.

This will be modified in the revised manuscript.

- l. 36: Sommer et al. (2018) focuses on the impact of wind on snow packing. There are many other papers more appropriate to refer to the impact of wind on snow cover variability.

We will look more carefully at the past literature in order to cite papers more focused on the impact of wind on snow cover variability.

- l. 40-52: The references of snowpack simulations driven by NWP data at kilometric resolution are the same as in Quéno et al. (2016). More recent publications should be included to update the literature review.

We will look more carefully at recent literature in order to update the literature review.

- l. 100: please define Cfs. I can't find what it means in the Köppen system, typo?

Yes, it is a typo, since the inner, mountain part of central Italy is Cfb. Also, we would add the "inner" attribute to "central italy" in the text, for a better geographical reference.
After the check, we also decided to add a further reference, since the Köppen classification was adapted by Pinna (1970) for Italy.

Pinna, Mario, Contributo alla classificazione del clima d'Italia, «Rivista geografica italiana» LXXVII, II, 6/1970, p. 129-152.

- l. 117-122: the presented snow climatological data is very old (1921-1960). Isn't there more recent climatological data over the region? i.e., more recent statistics, instead of simply mentioning a decrease in the past decades (l.122-123)?

A more detailed discussion about these aspects and more recent references will be added to this section of the manuscript.
We are aware that the snow climatological dataset is dated by now. However, reports and maps provided by the Ministry of Interior in the early 70s represent the latest regional-scale information about the snow depth for the whole Italy. After that period, the fragmentation of the National Hydrographic Service into twenty different regional offices (one per Italian region) led to the loss of homogeneity in data collection and statistics. This issue is well known by the Italian hydro-meteorological community, as stressed by Libertino et al. (2018); Rossi (2020) and Alberton (2021). While several studies are available in the Alps and northern Italy, central and southern Italy suffer from a poorly densed snow gauges network. The data scarcity did not encourage research in this area.
Nevertheless, since central Apennines are particularly prone to avalanche hazards, which has caused frequent casualties in the last few years, some authors have focused their attention on the occurrence of extreme snowfall events. The most cited study in this field is Piacentini et al. (2020), where records of recent extreme snowfalls are provided. The occurrence of extreme snowfall and high snow occurrence/accumulation variability of this part of Italy, is also highlighted in a recent study from Fazzini et al. (2021), published after our manuscript submission.
The decreasing trend in winter precipitation in Central Italy in recent decades is mentioned by several studies, such as Brunetti et al. (2000), Pavan et al. (2008), Longobardi and Villani (2010). Romano and Preziosi (2013), Appiotti et al. (2014) and Scorzini and Lopardi (2019). The last paper also presents a comprehensive summary of available studies in Italy.

The last available statistics, specifically carried out on snow gauges data, is provided for the Gran Sasso d'Italia massif, the highest peak of Apennines, and is provided by the Interregional Association for coordination and documentation of snow and avalanche problems (AINEVA), where a decreasing number snow days are recorded on altitudes below 1300 m a.s.l, in the 30-year period 1978-2007, even if events with intense snowfalls are increasing in some specific areas (Romeo and Fazzini, 2008).

References

Alberton M. (2021) Water Governance in Italy: From Fragmentation to Coherence Through Coordination Attempts. In: Turrini P., Massarutto A., Pertile M., de Carli A. (eds) Water Law, Policy and Economics in Italy. Global Issues in Water Policy, vol 28. Springer, Cham. https://doi.org/10.1007/978-3-030-69075-5_15

Appiotti F, Kržélj M, Russo A, Ferretti M, Bastianini M, Marincioni F (2014) A multidisciplinary study on the effects of climate change in the northern Adriatic Sea and the Marche region (central Italy). Reg Environ Chang 14(5):2007–2024. https://doi.org/10.1007/s10113- 013-0451-5

Brunetti M, Maugeri M, Nanni T (2000) Variations of temperatures and precipitation in Italy from 1866 to 1995. Theor Appl Climatol 65(3): 165–174. https://doi.org/10.1007/s007040070041

Fazzini, M.; Cordeschi, M.; Carabella, C.; Paglia, G.; Esposito, G.; Miccadei, E. Snow Avalanche Assessment in Mass Movement-Prone Areas: Results from Climate Extremization in Relationship with Environmental Risk Reduction in the Prati di Tivo Area (Gran Sasso Massif, Central Italy). *Land* **2021**, *10*, 1176. https://doi.org/10.3390/land10111176

Longobardi A, Villani P (2010) Trend analysis of annual and seasonal rainfall time series in the Mediterranean area. Int J Climatol 30(10): 1538–1546. https://doi.org/10.1002/joc.2001

Libertino A., Ganora D. and Pierluigi Claps: Technical note: Space–time analysis of rainfall extremes in Italy: clues from a reconciled dataset, Hydrol. Earth Syst. Sci., 22, 2705–2715, 2018

Pavan V, Tomozeiu R, Cacciamani C, Di Lorenzo M (2008) Daily precipitation observations over Emilia-Romagna: mean values and extremes. Int J Climatol 28(15):2065–2079. https://doi.org/10.1002/joc.1694

Romano E, Preziosi E (2013) Precipitation pattern analysis in the Tiber River basin (central Italy) using standardized indices. Int J Climatol 33(7):1781–1792. https://doi.org/10.1002/joc.3549

Romeo V. and Fazzini M. (2008), La neve in Appennino. Prime analisi su 30 anni di dati meteonivologici. In: Neve e Valanghe, n. 63 aprile 2008, AINEVA. https://issuu.com/aineva7/docs/nv63

Rossi G. (2020) Institutional Framework of Water Governance. In: Rossi G., Benedini M. (eds) Water Resources of Italy. World Water Resources, vol 5. Springer, Cham. https://doi.org/10.1007/978-3-030-36460-1_4

Scorzini A.R. and Lopardi M., Precipitation and temperature trends over central Italy (Abruzzo region): 1951-2012, Theoretical and Applied Climatology, 135,959-977, 2019, https://doi.org/10.1007/s00704-018-2427-3

- Figures 1 and 2 are probably not necessary, given the large number of figures. A short overview of winter 2018-2019 without figures may be enough.

The figures will be moved in the supplementary material.

- l. 188-190: "With exception of the first simulation (…) in the study area." This sentence is very unclear to me. Please clarify.

This sentence is related to lines 199-201. Our simulations are constituted by a series of 60 hours model runs. In each simulation, the atmosphere is initialized with NCEP data, while the soil and snowpack are restarted with the previous run, in order to avoid discontinuities in snowpack simulation. Initial conditions for snowpack and soil are taken from NCEP data only for the first run of this series.

The text will be rearranged to make this point more clear.

- l. 199: WRF-Noah model runs from 1 December 2018: is it initialized with no snow cover at that date? If so, is it realistic? It is mentioned later for Alpine3D but should be clarified here too.

We have used the snow cover provided from NCEP data. However, on 1 December 2018, a thin and patchy snowpack was present only at the highest elevations of our domain that are not covered by the measurement network used in this work.

This point will be clarified in the revised manuscript.

- l. 220-221: please clarify for what the neutral atmospheric stability conditions are chosen: turbulent fluxes?

The atmospheric stability affects the computation of the heat fluxes. We imposed neutral conditions because it provided the best results in terms of simulated snow height compared to other atmospheric stability conditions used in several sensitivity tests.

- l. 227-229: Are WRF output variables also corrected for elevation difference between model grid cell and station, when compared to measurements?

The meteorological variables are not corrected by elevation difference when compared to measurements. The reason is due to the fact that we want to show the difference between Noah and Alpine3D, when both models are forced with the same meteorological variables. Meteorological variables can't be corrected for Noah, since it is online coupled with Alpine3D, and thus they are not corrected for Alpine3D too. This means that it is more interesting to compare measured meteorological variables and uncorrected meteorological variables since the latter is used to drive the snow models. However, some statistics on the elevation difference between a model grid cell and station elevation will be added to supplementary material, since they are fundamental to assess how real topography is represented in the model.

- l. 231-235: MBE, MAE and R are rather common metrics, which don't need explicit definition here.

As the reviewer suggests, the error metrics won't be explicitly defined in the revised paper.

- l. 236-237: As for meteorological variables, what about the elevation difference between model grid cell and station?

As already said for the meteorological variables, some statistics on the elevation difference between model grid cell and snow station elevation will be added to the revised manuscript.

- l.239-245 and l. 253-265: The definition of ETS, HI_{rdm}, Table 3, as well as the definition of Jaccard index, MDHD and ASSD are the same as in Quéno et al. (2016). Given that the validation method is the same as that paper, I would make the definition of validation metrics much more concise and refer to Quéno et al. (2016).

As the reviewer suggests, the aforementioned metrics won't be explicitly defined, but referred to Quéno et al. (2016)

- l. 250: please specify the value of the threshold SWE W_{max}.

W_{max} varies locally according to MODIS land use classification, so that it is not a single value. A table with the correspondence between land use class and W_{max} will be added to supplementary material.

- l. 252: how is the threshold of 51% chosen?

Since we had to build binary maps from snow cover fraction maps we decided to use a common threshold for MODIS and snow cover models, thus we arbitrarily decided to use the threshold of 51% to discriminate between snow and no-snow.

- Figure 5: plots are not very informative because of the large number of dots. Consider perhaps smaller dots, or density plots?

Since the number of dots in each plot is really large, decreasing the dot size does not significantly improve the plot readability. Thus we will use density plots, as the reviewer suggests.

- Section 4.1: all evaluation metrics about meteorological variables are strongly impacted by elevation difference between simulations and observations. If no elevation correction is made, some statistics on elevation differences should be provided.

As already said in previous answers, statistics on elevation differences between a model grid cell and station elevation will be added to the supplementary material.

- l. 323-325: regression lines don't give much information. It would be more informative to describe tendencies for each model (e.g. underestimation of highest snow depth by this or that model, etc.).

As the reviewer suggests, we will further compare observed and simulated snow height also in terms of maximum overestimation and underestimation. We will extend the same approach to the snow height variation analysis.

- Figure 7: consider using smaller dots to make it easier to interpret.

As for Fig. 5, we will use density plots to make the figure more informative.

- l. 358-361: "it is likely due to the more complex multi-layer Lagrangian scheme of the Alpine3D model (…)". There is no result in this paper to support this hypothesis. The effect of the different complexity of models on the snowpack evolution would be very interesting to compare, but there is no proof here to: 1) differentiate between snow compaction and melting; 2) claim that the more complex model reproduces better either compaction or melting.

The introduced snow density measurements gave the opportunity to calculate snow water equivalent. This allowed us to differentiate the contribution of melting and compaction to the observed snow settlement and compare it to model simulations.

- L. 369-374: Why not masking out forested areas when comparing MODIS to models?

Jaccard Index and ASSD can be computed only from binary matrices, so filled entirely by 0/1 or False/True values. Masking out forested areas, thus assigning to those cells an arbitrary value (-9999 on NaN for example) will make it impossible to compute the two indices. However the masking of the forested areas can be introduced for the comparison of the snow cover area fraction (Tab. 6 and Fig. 12), but then some metrics would be computed neglecting forested areas (MAE, MBE and R) while some other not (J and ASSD). In order to maintain coherence within the analysis, the authors decided to not mask out forested areas.

- Figure 10 could be included in Figure 9 to compare visually more easily and see more directly the impact of forested areas.

As the reviewer suggests, the forested areas of Fig. 10 will be superimposed on Fig. 9 in the revised manuscript.

- l. 392 and Table 6: probably a factor 100 error (rather 6% than 0.06%)

This is not an error. Snow cover area fraction ranges from 0 to 0.3%, thus MBE and MAE equal to 0.06% are reasonable values.

- Section 4.2.2: as far as I understand, the snow cover fraction is computed a posteriori using equation (6). However, the fractional snow cover can have an impact on the snow cover dynamics, especially at 3 km resolution. Why not using a fractional snow cover scheme in the models, e.g. Helbig et al. (2021) or simpler schemes?

Noah implements Eq. 6 thus it calculates at runtime the snow cover fraction. To the authors knowledge, Alpine3D does not implement schemes for snow cover fraction calculation, thus the authors applied Eq. 6 to Alpine3D a posteriori in order to derive snow cover fraction also for that model and compare it to MODIS. Helbig et al. (2021) scheme will be taken into account for future works on this topic.

- l. 416-417: use North-South instead of upper part-lower part which could suggest elevations.

We will apply the suggested correction in the revised manuscript.

- l. 431-432: accumulation should be used for snow water equivalent increase. Please reformulate the sentence, and in particular "snow height accumulation" and "snow water equivalent accumulation".

As the reviewer suggests, “snow height accumulation” and “snow water equivalent accumulation” are not correct in this sentence. We will reformulate the sentence in the revised manuscript, also highlighting the differences with the measured snow water equivalent.

- l. 428-443: The interpretation of Figure 15 is a bit confusing. Apart from a little more SWE in Noah than Alpine3D in early December, the SWE is always higher for Alpine3D than Noah. Furthermore, a more in-depth analysis separating snow accumulation and melting amount would be necessary to draw more solid conclusions which would allow to better clarify the differences between models.

As already said in previous answers, the access to measured snow height and density data will permit us to compute measured snow water equivalent. Thus it will be possible to compare measured and modeled snow water equivalent and draw more solid conclusions on the differences between the models.

- Conclusion: values from the results should not be repeated there but summarized.

The suggested correction will be applied to the revised manuscript.

- Conclusions linking the bulk performance metrics to the different representation of processes in the models are not proved, only hypotheses at this stage.

We will put more effort to identify the main causes of differences between the models, also considering snow water equivalent.

Local remarks and typos

- l.9: define LSM
- l. 31: typo snowpack
- l. 35: precipitation, singular
- l. 60: snow layers
- l. 104: altitudes -> elevations
- l. 137, l. 142: leaded -> led
- l. 148: performance
- l. 154: delete area or domain.
- l. 154: near-surface
- l. 167-170: to be deleted, redundant.
- l. 194: a model
- l. 281: altitude -> elevation
- l. 313: the three time series
- l. 320: most of the 13 sites
- l. 436: It is
- l. 450: comparing them to
- Language proofreading by a native speaker could be useful.

These typos and remarks have been all corrected.

References

Essery, R., Morin, S., Lejeune, Y., and Menard, C. B.: A comparison of 1701 snow models using observations from an alpine site, Adv. Water Resour., 55, 131–148, doi:10.1016/j.advwatres.2012.07.013, 2013.

Helbig, N., Schirmer, M., Magnusson, J., Mäder, F., van Herwijnen, A., Quéno, L., Bühler, Y., Deems, J. S., and Gascoin, S.: A seasonal algorithm of the snow-covered area fraction for mountainous terrain, The Cryosphere, 15, 4607–4624, https://doi.org/10.5194/tc-15-4607-2021, 2021.

Quéno, L., Vionnet, V., Dombrowski-Etchevers, I., Lafaysse, M., Dumont, M., and Karbou, F.: Snowpack modelling in the Pyrenees driven by kilometric-resolution meteorological forecasts, The Cryosphere, 10, 1571–1589, https://doi.org/10.5194/tc-10-1571-2016, 2016.

Citation: https://doi.org/10.5194/tc-2021-285-RC1

REVIEW 2

Review of "Snow cover prediction in the Italian Central Apennines using weather forecast and snowpack numerical models", by Raparelli et al.

In "Snow cover prediction in the Italian Central Apennines using weather forecast and snowpack numerical models" the authors present an application of the Noah LSM and Alpine3D snowmodels forced with WRF 3km data. Overall this manuscript is well written and generally clear. Figures are mostly clear and I enjoyed the contextualization of the climate during the study period.

My main concern is that at the 3km lateral spatial resolution, we know a point-scale model is not representative of the sub-grid dynamics across such a large area -- blowing snow, canopy interactions, avalanching, sublimation on exposed ridges, variable compaction rates, surface exchange, etc. The authors note this but I'd like to see more attention drawn to it as a (major in my opinion) limitation. Indeed I would like to see more discussion on this scale mismatch between the vertical resolution -- O(mm) and O(cm) scales with SNOWPACK -- and the lateral O(km) scales. Reposed, is the computational effort of running Alpine3D warranted over a simpler snow model but at a finer spatial resolution? What are the process representations in SNOWPACK that have the largest impact on the simulations versus Noah? As there seems to be no treatment for sub-grid fractional snow cover, it isn't clear to me how the vegetation interactions are considered in this context. Lastly, I'd like to see a bit more detail regarding how wind downscaling was done as it wasn't completely clear -- is it straight out of WRF using the 3km topography? This is a large source of uncertainty with respect to surface exchanges.

In general I'm left a bit uncertain what the novel take away of this manuscript is. The application domain is certainly unique but I'm left feeling that there is a missing discussion on the "why" of comparing these two models. Specifically, how the process representation in these models and the uncertainty in parameter estimation leads to "better" or "worse" process representation. Figure 15 suggests some very fundamental differences outside of the snow microsturcture scheme resulting in such large SWE differences. I'm missing a deep understanding as to what is causing this.
I would suggest moderate revisions to contextualize the 'why' better.

We are very thankful to the reviewer for the time spent evaluating our work and we thank him/her for the useful and constructive comments. We have answered all comments. Note that our answers are in **BLUE** in the following text.

Comments to reviewer's general comments:

"Is the computational effort of running Alpine3D warranted over a simpler snow model but at a finer spatial resolution?"
In this paper we wanted to investigate the performances of a simple and a more complex model in simulating snow height and snow cover extent on a regional scale at kilometric horizontal resolution. We chose as study domain Central Apennines, a territory characterized by complex topography where no one conducted this kind of study before. We showed that already looking at snow height and snow cover extent the models show different performances, but to increase the robustness of the analysis we will include in the revised manuscript also the comparison of simulated and measured snow density and snow water equivalent data. These further comparisons will show more evident differences between the model and it will help to conclude if it makes sense to run Alpine3D at 3km resolution for the estimation of snow height, snow density, snow water equivalent and snow

extent, when already Noah simulations are available. However, we also think that thanks to the implementation in Alpine3D of more detailed physics processes, the model will benefit more than Noah in terms of performances if the spatial resolution of the meteorological forcing is increased, but this is an open question and it will be investigated in future works.

“What are the process representations in SNOWPACK that have the largest impact on the simulations versus Noah?”
We believe that the comparison of measured and simulated snow density and snow water equivalent will help to better understand which are the processes that cause the largest differences between the models. This aspect will be further investigated in the revised manuscript.

“As there seems to be no treatment for sub-grid fractional snow cover, it isn't clear to me how the vegetation interactions are considered in this context.”
Actually, Noah already takes into account vegetation and computes fractional snow cover. The snow-canopy interaction was instead disabled in Alpine3D, but it will be activated in the simulations that we will show in the revised manuscript. Indeed we also want to extend the simulations to winter seasons 2019-2020 and 2020-2021, for a total of 3 simulated seasons, in order to take into account winter interannual variability and increase the robustness of the analysis.

“Lastly, I'd like to see a bit more detail regarding how wind downscaling was done as it wasn't completely clear -- is it straight out of WRF using the 3km topography?”
We didn’t downscaled wind, and we didn’t downscaled other variables. We just regridded on regular grid-spacing at a resolution of 0.03°, very close to the native horizontal grid adopted in WRF (about 3 km), which is curvilinear and not supported by Alpine3D.

Specific comments:

Abstract:
Can be tightened up with the results more succinctly summarized
The abstract will be summarized in the revised manuscript.

L8 "online" -> Is this a two-way coupling?
With online coupling, we mean that WRF drives Noah, which gives feedback to WRF. This is also called two-way coupling, as the reviewer suggests.

L 9 "LSM" remove.
This was a typo error and it will be removed.

L 9 "Alpine3D"
As I understand it Alpine3D is essentially met downscaling/interp + blowing snow + distributed SNOWPACK. I think this should clearly stated as to what the underlying snowmodel is as that is most relevant to this study.
Alpine3D is a snowpack and soil numerical model. It includes models for snow cover (SNOWPACK), vegetation and soil, snow transport, radiation transfer, and runoff which can be enabled or disabled on demand. In our setup Alpine3D is forced with the 1-hour meteorological data taken from WRF outputs. These data were not downscaled, but they were only regridded on regular grid-spacing at a resolution of 0.03°, very close to the native horizontal grid adopted in WRF (about 3 km), which is curvilinear and not supported by Alpine3D. Moreover, the aeolian snow transport and radiation transfer modules were not activated in our Alpine3D configuration. This point will be clarified in the revised manuscript.

L23 "at moderate spatial resolution (3 km)"
Be explicit that the snowmodel was run at this resolution.
It will be explicitly stated in the revised manuscript

L31 snowpacl -> spelling
It will be corrected in the revised manuscript

L35 precipitations -> No s
It will be corrected in the revised manuscript

L55 Suggest you include Vionnet, 2021 (https://doi.org/10.5194/tc-15-743-2021)
We were not aware of this paper, and we thank the reviewer for the suggestion. The paper will be included in the revised manuscript

L100 "Classified as Cfs"
Please remind the read what Cfs stands for.
It is actually a typo, since the inner, mountain part of central Italy is Cfb, which means temperate oceanic climate (C="temperate", f="no dry season", b="Warm summer"). It will be better specified in the revised manuscript. The Cfs is not actually an acronym, but a standard climate classification proposed by Koppen (referenced in the manuscript)

L103 "thermal excursion"
is not clear to me in this context. word choice or can you clarify?
It refers to the difference between the minimum and the maximum temperature value observed in one day (24h). To better clarify it is referred to the air temperature, we change it with 2temperature excursion"

L111 "being the yearly maxima mainly localized in the western slope"
This is not clear
The Apennine mountain chain crosses the italian peninsula from north-west to south-east, facing the Adriatic sea at the east side and the Tyrrhenian sea at the west side. It is the west side of the Apennines that experiences the maximum yearly precipitation. It will be clarified also in the revised manuscript.

L114 "it is worth to mention"
-> It is worth mentioning
although this is a bit colloquial. I suggest you change.
It will be corrected in the revised manuscript

L117 "the same"
The same as what?
The same Apennine area described in the previous paragraph (Central Apennines). We will better clarify in the revised manuscript

L118 "on annual basis"
On /an/ annual basis
It will be corrected in the revised manuscript

L135 "the second decade"
Decade? Do you mean "the second half"?
The sentence will be rewritten as: Between 10th and 20th December…

L137 "small cold"
Small, cold
It will be corrected in the revised manuscript

L 137 "impulse"
Uncertain of meaning in this context. Word choice?
The sentence will be rewritten as: In the middle of December, cold air advection from Northern and Northern-Eastern europe leads …

L137 lead~~ed~~
It will be corrected in the revised manuscript

L140 "determined"
Caused?
The revised manuscript will be corrected according to the reviewer's suggestion

L 142 "impulse"
Uncertain of meaning in this context. Word choice?
The sentence will be rewritten as: when cold air advection from North Atlantic…

L152 ~~automatic weather stations~~ and use previously defined AWS only
It will be corrected in the revised manuscript

L154 "The measured variables"
List units and how frequently they were visited please
The civil protection weather station network includes automatic stations which measures surface air temperature (°C), relative humidity (%), wind speed (m/s), incoming shortwave radiation (W/m2), precipitation (mm/1h) and snow height (cm). Data acquisition is automatic and all measurements are transmitted every 15 minutes through GMS or radio links to a central server, thus the data were not collected manually. That information will be specified in the paragraph.

L190 "for WRF,"
remove comma
It will be corrected in the revised manuscript

L190 "it resulted the best"
Resulted in
It will be corrected in the revised manuscript

L199 "is run from 1 December 2018"
Why not the commonly used October 1? Can the authors confirm there were no missed snowfall events? How does this impact soil temperature?
We confirm that at automatic weather station locations the first snowfall occurred around mid-December, thus we didn't miss previous snowfall at weather station sites. Moreover, the soil conditions were initialized with NCEP reanalysis and afterwards driven by WRF simulations. Thus the first snowfall wasn't right after the soil initialization.

L220 "After a sensitivity test (not shown here)"
What forcing variables/parameters were tested? How was this done?
The sensitivity test was conducted using several combinations of different parameterizations available in Alpine3D. A more detailed description on the methods used for the sensitivity test and the parameters involved will be added to the supplementary material

L221 "as well as the selection of the "Zwart" and "Lehning new""
Please describe what these are including the associated references. How do these compare to what was selected in Noah?

Actually, the parametrization used was “Lehning_1” and not “Lehning new”, However, to the author's knowledge, there are no papers where these parameterizations are explicitly described. They can be found in SNOWPACK/Alpine3D source code, and will be explicitly described in the supplementary material

L224 "of the simulation results"
Of what, specifically?
In terms of snow cover runoff. It will be specified in the revised manuscript

L417 "lower part"
Lower elevation?
With lower part we ment southern part. Il will be corrected in the revised manuscript.

Figures:
All timeseries figures need a starting x=0 tick and in general all would benefit from more x-axis ticks
The reviewer suggestions will be applied in the revised manuscript.