Dear Jakob,

We are happy to resubmit a revised version of the manuscript "A dynamical and thermodynamic mechanism to explain heavy snowfalls in current and future climate over Italy during cold spells" now entitled "Present and future synoptic circulation patterns associated with cold and snowy spells over Italy". We understand that the reviewers and the editorial board have raised several issues with our manuscript and we have undertaken the suggested changes. This resulted in additional work which Miriam D'Errico could not completely afford due to her new job in education and few personal unfortunate events. We have therefore asked Flavio Pons, who is currently working in closely related topics on cold spells over France, to take the lead of the analyses and produce a new version coherent with the reviewers and editorial comments. All co-authors agree that Flavio and Miriam will share the first author position of the paper.

About the paper, we would like to stress that:

1) **It is now oriented to describing the dynamics of cold spells events, namely the associated circulation patterns and their changes. As demanded by the reviewers, we leave aside the thermodynamic changes that cannot be proficiently investigated using PlaSim.**
2) **We defend the choice of using PLASIM for circulation studies with the possibility of performing very long stationary climate simulations and to obtain good analogues of atmospheric circulations associated with cold spells. To this purpose, we also perform a bias correction which ensures that the control run has analogues statistics of temperature patterns compatible with those extracted from NCEP.**
3) **Since the reviewers questioned our claim that "the detected [cold-spell] events have a common dynamical signature", we have performed a cluster analysis of circulation patterns. This analysis clearly shows that those patterns can be divided into two clusters, of which we provide extensive analyses in the new version of the manuscript.**

We hope that the proposed changes constitute a sufficient improvement to meet ESD publication standards. We would be very thankful if the previous reviewers, which provided useful insights and comments on our work, would be willing to review the current version of the manuscript. Below, we provide an updated answer to the major comments of the reviewers, ensuring that we have taken into account the minor remarks while writing the new version of the manuscript.

Best Wishes,
Davide (on behalf of all the authors)

Reviewer 1

GENERAL COMMENTS

This paper first uses reanalysis data to document the large-scale circulation conditions that have led to heavy snowfall events together with low temperatures in Italy between the years 1954 and 2018. After this, an intermediate complexity climate model (PlaSim)is used to explore how the occurrence of such events might change in a warmer future climate.The paper has its positive aspects but also severe limitations. To start with the former,it includes a valuable compilation of 32 major snowfall / cold spell events that have affected Italy since the mid-20th century. The analysis of the corresponding circulation anomalies in the NCEP reanalysis also makes good sense. Furthermore, the paper is written in good English.

On the other hand, there are many problems in the PlaSim simulations and in the interpretation of their results. The first is, obviously, the coarse (2.8◦×2.8◦) resolution of the model. Simulations at such a resolution give very little direct information on snowfall in Italy. In the control simulation, the (country mean) snow water equivalent anomalies in the identified extreme cases are of the order of 2 mm, which is at least an order of magnitude smaller than the observed local snowfalls. Therefore, in practice, the analysis mainly gives information on the atmospheric circulation events that resemble the circulation during the observed snowfall extremes.

We thank the reviewer for this comment. We acknowledge that, in the previous version of the manuscript, we have used PlaSim beyond its capabilities to study small scales thermodynamic processes. Indeed, we largely reformulated the paper giving much more focus to the large scale circulation than to the phenomena at the ground.

A second important problem is that the severity of the cold spells is only analysed based on the magnitude of the 850 hPa cold anomalies relative to the climatological mean values. These anomalies are found to (more or less) retain their earlier magni-tude, leading to the suggestion that such events in the future will be no less severe than those observed this far. Yet the warmer mean climate in these simulations also means that the actual temperatures during the cold spells will become higher. In the RCP8.5scenario, this change is large enough to nearly eliminate all snowfall in Italy. Thus, a cold anomaly with the same magnitude will not have the same effects in a warmer climate.

In the new version of the paper, we rely on absolute temperature rather than anomalies. We have also performed bias corrections against reanalysis data to ensure coherence between PlaSim and NCEP.

In addition to the simulation based on the RCP8.5 forcing scenario, the study uses another simulation in which the sea surface temperature (SST) has been uniformly increased by 4 K, without changing the atmospheric composition. Such a simulation may be useful for process studies but does not represent a plausible future. Increasing the SST without increasing atmospheric greenhouse gas concentrations creates an artificial energy source at the sea surface, which distorts the dynamics of the climate system. The finding that the simulated snowfall extremes increase under such conditions is therefore difficult to interpret.

We do agree with the reviewer that the +4K SST is an unphysical scenario. This is now clearly stressed in the new version of the manuscript. When using this scenario our objective is to analyse the retroaction of high SST values on the atmospheric circulation. Indeed in the area of study (Mediterranean Basin), the SST are expected to increase faster than in the surrounding oceans. We believe that this scenario is still useful because we show that warmer SST do not produce any major change in large scale dynamics, with results comparable with the control run.

Aside from these scientific issues, the selection of figures requires consideration. For Example, Figure 8 is hardly at all discussed in the text, suggesting that it is redundant. Figure 4 is also a candidate for deletion (see comment 17 below). On the other hand,to aid the reader to assess how severe the simulated future cold spells are, figures and/or other information on the average winter warming would be needed.In conclusion, large improvements are still needed in this paper

The structure of the paper has been changed radically and only Fig. 1 was maintained from the previous version. We hope that the reviewer will appreciate the new insights of the new version of the manuscript

Reviewer 2

The paper discusses the topic of heavy and extreme snowfall in Italy in current and future climate. This is a scientifically interesting and societally relevant subject. The Starting point of the analysis is a set of 32 extreme historic cases with significant snow-falls in at least one of two Italian cities of Bologna and Campobasso. The author's goat length in describing these cases (in the Appendix), which vary from relatively short outbursts, to long-lasting episodes involving cold spells in large parts of Europe. This is followed by an analysis of snowfall under similar circulation types, occurring in 500 year simulations conducted with an intermediate complexity model (PlaSim). It is concluded that extreme snowfall may increase or decrease, depending on whether or not future climate change will express more than average warming of the Mediterranean.The paper provides an interesting set of observed cases, along with some interesting analysis of simulations in a coarse resolution intermediate complexity model. How-ever, as the paper is presently formulated, it lacks to provide a convincing story that connects the two. There are moreover serious shortcomings in the current description and presentation of the results, which I will try to motive in more detail below. Based on this, however, I recommend to reject the paper in its current form.

Major remarks

1. Event definition. In the Introduction the authors argue, that while there is general consensus that temperature is increasing and mean snowfall is decreasing, knowledge of the changes of extreme “snowy” cold spells is inconclusive, because of inconsisten-cies in their respective definitions. From this statement I had anticipated that the paper would start with such a definition. However, it is absent. Instead the authors implicitly“define” the case by means of the observed large-scale circulation that accompanied the (start of the) events. Despite the circulations being “very similar” as the authors write on p5 L140, there is apparently enough variation to allow the huge differences in the observed snowfall amount (Fig3). The correlation figures, though only briefly described, also seem to hint in this direction (rather low correlations).

In the new version of the paper, we classify the 32 events based on the baric configuration, showing that they can be divided into two clusters, where the synoptic situation corresponds to an omega blocking with high pressure on the UK, and to a Scandinavian positioning of the high pressure center, with Easterly flow towards the Mediterranean. Then, the search for similar events in the climate simulations is carried out using analogues based on the dynamic configuration of each event. The high phenomenological variability observed in the events remains; however, the 32 events were indeed chosen because they caused significant impacts at sub-country level in Italy. Using dynamic analogues, we search for simulated events that have the same potential to produce similar situations due to the synoptic configuration of the pressure field.

2. Snowfall/depth in intermediate complexity models. The way in which the study at-tempts to address its main question, involves the use of an intermediate complexity model. While there is nothing wrong with using such intermediate complexity models,it can be questioned whether they are suitable for the problem at hand. Cold spells, es-pecially when defined with respect to a fixed temperature, and in particular snowfall,will depend sensitively on a lot of parameters, microphysics, precipitation, the repre-sentation of the underlying orography and much more. Since for snowfall to occur, the temperature has to be around freezing point, biases in temperature will all too easily imply biases in snowfall. To the knowledge of this reviewer, intermediate complexity models are relevant to the real world mostly because of their reasonably well resolved“dynamics", not so much because of the details of their resolved thermodynamics /C2 Surface parameters / precipitation, let alone snowfall. As a consequence, I think the results in this paper should be treated with extreme care, and can basically only be interpreted within the limited validity of the intermediate complexity model itself, and not as a direct proxy of what may happen in the real world at a local scale, such as, in this case, in Italy.

We have completely rephrased the manuscript and presented our approach as a possible way to detect compound extreme events from documentary sources and then analyse the role of atmospheric circulation in their occurrence via a simple model capable of producing several hundred thousands of large scale circulation

patterns in relatively short time. We will take particular care in clearly stating the thermodynamic limitations of our study.

3. Reanalysis.The principal source of reanalysis data is well known for its shortcomings, of especially its surface variables.Some reasons aregiven in https://journals.ametsoc.org/bams/article/77/3/437/55258/The-NCEP-NCAR-40-Year-Reanalysis-Project. As such it is questioned whether the snowfall, t2m tem-perature and consequently snow depth are variables that can be meaningfully used.Upper-level air temperature, and Z500, as well as possibly mean sea-level pressure can be safely used.

In the new formulation of the paper, we base our analysis on sea level pressure and 850 hPa temperature for the characterization of the events and the definition of the analogues, and we leave the analysis of snow depth more as an additional comment, without it being a crucial parameter in the definition of PlaSim coldspell events.

4. Unrealistic SST+4K simulation Three different simulations are carried out with PlaSm. In one of them the global SST is increased uniformly by 4 degrees. By not changing atmospheric forcing, this leads to an unrealistic situation. The situation of lakeside snow effects might be an important aspect of snowfall changes in the future,but it is likely that some sort of compensating effect occurs in reality. As a consequence, the statements in this paper are likely over-confident. Without doubt there is arole for both circulation and thermodynamic processes. It is worthwhile to look up some recent literature by e.g. O'Gorman on this subject.

We do agree with the reviewer that the +4K SST is an unphysical scenario. This is now clearly stressed in the new version of the manuscript. When using this scenario our objective is to analyse the retroaction of high SST values on the atmospheric circulation. Indeed in the area of study (Mediterranean Basin), the SST are expected to increase faster than in the surrounding oceans. We believe that this scenario is still useful because we show that warmer SST do not produce any major change in large scale dynamics, with results comparable with the control run.

5. Statistical significance. The study starts with a description of the 32 cases (or in fact the description is only given in the appendix). Reading through this interesting and expansive list I get the conclusion that there is a substantial difference between the historic cases, both in scale, in duration, in extremity, etc. As exemplified by Fig3 the variance in local snowfall accompanying these events is huge. Despite this variance,the authors state that the underlying T850/SLP or Z500/SLP fields are quite similar.Why then, do the users restrict themselves to use only 32 cases from the simulations?To me this is unclear. It basically means that for every historic event, only the closestC3single model event is selected, whereas already from the observations it becomes clear that there is a huge variability within these cases. In other words, there must be many similar

circulations where no snowfall occurs. I could imagine that more robust (model)results could be obtained by considering a larger subset of similar circulation types.

In the new version of the paper we have radically changed the design of the numerical experiment. Now we look for analogues being the closest x% of the PlaSim simulated field to the event of interest, leading to more robust samples.

6. Given my comments above, it is my feeling that the paper could benefit from a radical change of viewpoint. By letting the simulations of the model of intermediate complexity from the heart of the paper, and providing context from observed cases in an added discussion, the claims could be made more specific to what is achievable with such a model. For example, how do cold spells change in such a model, and can these be used to examine extreme snowfall. Because you run a simple model, you can afford to run as many long simulations as are required to achieve at least significant results with respect to the circulation changes. The thermodynamic changes will be hard given the limitations of the model, but perhaps some knowledge can be squeezed out, if results are considered at larger spatial scales. I do not think PlaSim can be reasonably expected to give realistic results at local scale.

Following the suggestion of the reviewers, we have completely reorganized the manuscript. First, we have given more space to the reconstruction of the 32 events, which was one of the major challenges of this study. Then, we have specified the use of intermediate complexity climate simulations as a tool to study the atmospheric circulation associated with the detected extreme events, through the use of analogues. Finally, the +4K simulation has been exploited only as an additional tool to understand the possible role of a warmer mediterranean sea in determining a thermodynamic feedback to the atmospheric circulation associated with these events. We have underlined the limitations and suggest our approach (documentary sources + analogues search) as a way to investigate compound extremes.