We thank the reviewer for the careful reading of the manuscript and the constructive comments. Below we provide a point-by-point response.

Comment 1:

Motivations and originality. In the Introduction, you note that several other deterministic reanalyses/hindcasts already exist over the same domain, at resolutions that are not substantially different from that of MORE. Yet in the Conclusions, you characterise MORE as a major achievement. Intuitively, I would associate a truly major breakthrough with a product offering a qualitatively new capability e.g. such as a convection-permitting ensemble hindcast, to give one example. This raises a question that I believe deserves more explicit attention: what motivated the development of yet another deterministic hindcast over this domain?

We thank the reviewer for this important comment. We agree that MORE should not be presented as a fundamentally new modelling capability or as a replacement for existing regional reanalyses and hindcasts. Rather, its value lies in providing an additional independently generated convection-permitting dataset over Italy and the Alpine region, based on a different modelling framework and downscaling strategy than existing products.The motivation for developing MORE is multifaceted. Independent high-resolution datasets are valuable for assessing uncertainties and improving the robustness of meteorological and climatological analyses through multi-dataset intercomparison. In addition, MORE provides a long-term convection-permitting dataset at spatial scales below 2 km, specifically designed for applications over Italy and the Alpine region.

A further distinctive aspect of MORE is its comprehensive multivariable and three-dimensional structure. Beyond near-surface variables, the dataset includes atmospheric fields such as geopotential height, temperature, humidity, and wind components at multiple pressure levels, enabling physically consistent analyses across different spatial and vertical scales. This characteristic considerably broadens the range of potential applications of the dataset, including process-oriented studies, hydro-meteorological applications, and emerging artificial intelligence methodologies that require long-term, high-resolution, physically consistent atmospheric datasets.

We agree that these aspects were not sufficiently emphasized in the original manuscript and will revise the Introduction and Conclusions accordingly.

Comment 2:

Verification scope. There is a clear imbalance between the number of variables produced and delivered by MORE and the two variables selected for verification. I acknowledge that soil properties are implicitly evaluated through the application sections. In general, I agree that 2-metre temperature (t2m) and total precipitation (tp) are likely to be the most widely used variables among the target user community, and that they benefit from relatively dense observational networks, making observation-based verification straightforward. However, in its current form, the manuscript does not clearly articulate the rationale behind this choice, nor does it explain why upper-level atmospheric variables are excluded from the verification exercise. A comparison between ERA5 and MORE at upper levels would, for instance, have been an informative addition. I am not asking the authors to perform such additional analysis, but rather to provide a transparent justification, either in the Introduction or in the

Methods/Evaluation section, for the decision to focus exclusively on surface variables. Where relevant, the authors could also direct readers to existing literature in which MOLOCH has been evaluated against other variables (e.g. wind, pressure, upper-level temperature), thereby contextualising the verification choices made here and reassuring readers that the model's performance for those variables has been assessed elsewhere.

We thank the reviewer for this thoughtful comment. We agree that the manuscript does not currently provide a sufficiently explicit justification for focusing the evaluation on precipitation and near-surface temperature, despite the much broader range of variables available within MORE.

The choice was primarily motivated by the fact that precipitation and 2-m temperature are among the most widely used variables for hydro-meteorological, climatological, and impact-oriented applications. In addition, they benefit from dense and long-term observational datasets over Italy, allowing a robust multi-decadal validation framework based on independent observations.

We also agree that MORE includes many additional variables, including upper-air diagnostics, wind fields, soil variables, and radiation-related quantities, which are potentially of great interest to users. However, a systematic validation of the full set of variables available in MORE would require dedicated analyses that go beyond the scope of the present manuscript, whose primary objective is to introduce the dataset and provide a first comprehensive assessment of its performance.

Following the reviewer's suggestion, we will revise the manuscript to better clarify the rationale behind the verification strategy adopted in this work. We will also provide references to previous studies in which the MOLOCH model has been evaluated for additional atmospheric variables and upper-level fields, thereby offering readers further context regarding model performance beyond precipitation and near-surface temperature.

Comment 3:

Communication of evaluation results. I have some uncertainty regarding the overarching motivation behind the evaluation as presented. In places, the text and its interpretations read as a sanity check of the MOLOCH downscaling. This is a legitimate goal, but the suitability of MOLOCH for dynamical downscaling is already well supported by prior literature. I would argue that the most important message of the evaluation is currently underemphasised, and that the text would benefit from revision to remove repetitions (e.g. the concepts in lines 711-714 are repeated several times across the manuscript) and improve readability.

In my view, the central contribution of the evaluation should be to give prospective users a concrete sense of the accuracy and precision of MORE for the verified variables at the temporal aggregations considered. Specifically, I would encourage the authors to address the following questions explicitly: What is the typical uncertainty of daily precipitation estimates from MORE, expressed, for instance, as a relative error? How does this change at hourly aggregation? What uncertainty should a user expect when working with 2-metre temperature? The answers to these questions are arguably the most practically valuable outcome of the evaluation, yet they are not sufficiently prominent in the current manuscript.

Equally important is guidance on the appropriate use of MORE: for which types of analysis is the product recommended? Are there specific regions, seasons, weather regimes, or applications for which MORE is known to perform less reliably, and where its use should therefore be approached with caution? This information, both the quantified uncertainty estimates and the

user guidance, should be given a much more prominent role in the paper, and should be reflected clearly in the Conclusions and the Abstract.

We thank the reviewer for this important and constructive comment. We agree that one of the primary objectives of the evaluation should be to provide prospective users with a clear understanding of the expected accuracy, uncertainty, and limitations of MORE for the variables and temporal scales considered.

While part of this information is already contained in the current evaluation, we acknowledge that it is not sufficiently emphasized and that some discussions focus excessively on intercomparison aspects. We agree that a more explicit interpretation of the results would substantially improve the practical value of the manuscript for future users of the dataset.

In the revised manuscript, we will therefore strengthen the discussion of the expected uncertainty ranges associated with precipitation and near-surface temperature, highlighting how these uncertainties vary across temporal scales, seasons, and geographical regions. We will also provide clearer guidance regarding the appropriate use of MORE, including a discussion of the main strengths and limitations emerging from the validation and application analyses.

More generally, we will revise parts of the evaluation section to reduce repetitions and improve readability, placing greater emphasis on the information that is most relevant for prospective users, namely the expected accuracy of the dataset, the interpretation of its uncertainties, and the contexts in which its use is most appropriate.

Comment 4:

Lines 113-119. The statement on "ensemble approaches" is not clear. Consider moving this paragraph to Conclusions?

We thank the reviewer for this comment. We agree that the original use of the term "ensemble approaches" was potentially ambiguous and could be interpreted as referring to ensemble forecasting methodologies, which was not our intention.

Our objective was instead to emphasize the value of MORE as an independent convection-permitting dataset contributing to multi-dataset analyses and intercomparison frameworks over Italy and the Alpine region. We agree that this aspect was not sufficiently clear in the original manuscript and we will revise the corresponding paragraph accordingly, clarifying the role of MORE within the growing set of high-resolution regional reanalyses and hindcasts and avoiding the ambiguous use of the term "ensemble approaches".

Comment 5 :

Lines 211 and following paragraphs. The manuscript would benefit from a more explicit statement of the objectives underlying the two distinct strands of the evaluation. What insights do the authors expect to gain from comparing MORE against observations? And what is the purpose of comparing MORE against other reanalyses? I believe the answers are implicit in the text, but leaving them for the reader to infer is an unnecessary source of ambiguity. Once again, I would encourage the authors to state these motivations clearly and upfront. For instance, framing the comparison against observations as an assessment of the accuracy and precision of MORE, and framing the comparison against other reanalyses as a sanity check confirming that

the downscaling procedure does not introduce gross inconsistencies relative to established products.
On a related point, there is a degree of internal tension in the manuscript: the authors state that they do not intend to rank the reanalyses or hindcasts (line 308) under comparison, yet the structure and language of the evaluation section largely does exactly that. I would ask the authors to resolve this inconsistency. If the comparison with other reanalyses is intended as a sanity check rather than a performance ranking, then the criteria for a satisfactory outcome should be made explicit. What would a successful sanity check look like? What level of agreement, or disagreement, with other reanalyses would the authors consider acceptable, and on what basis? Providing clear answers to these questions would considerably improve the interpretability of this section for the reader.

We thank the reviewer for this important and insightful comment. We agree that the original manuscript does not sufficiently distinguish the different objectives underlying the comparison against observations and the comparison against other reanalyses and hindcasts, which may lead to ambiguity in the interpretation of the evaluation framework.

Our intention is indeed twofold. First, the comparison against observations is aimed at assessing the accuracy, uncertainty, and realism of MORE in reproducing precipitation and near-surface temperature variability across different temporal scales and climatic conditions. Second, the comparison against existing reanalyses and hindcasts is intended to place MORE within the context of currently available high-resolution products and to assess whether the long-term convection-permitting downscaling strategy produces physically consistent behaviour relative to established datasets.

We also agree that the manuscript currently contains some tension between the stated objective of avoiding a formal ranking of products and parts of the discussion that may inadvertently be interpreted in that way. Our intention is not to identify a “best-performing” dataset, but rather to evaluate the consistency of MORE with observations and with the range of behaviours exhibited by existing high-resolution products, while also assessing the capability of kilometre-scale convection-permitting downscaling to realistically represent fine-scale atmospheric variability.

Following the reviewer’s suggestion, we will revise the evaluation framework to make these objectives more explicit, clarify the interpretation of the intercomparison, and reduce ranking-oriented language throughout the manuscript. We will also better explain what is meant by a satisfactory level of consistency between MORE and existing products, and how this information should be interpreted by prospective users of the dataset.

Comment 6 :

If the comparison with other reanalyses is intended as a sanity check, upscaling MORE to the coarser reference grid would be a more methodologically appropriate basis for the intercomparison. Evaluating global reanalyses on the native high-resolution grid of a convection-permitting product places them in an inherently unfavourable position (lines 398-402). That said, such a comparison may carry practical value by illustrating quantitatively what is lost by forgoing dynamical downscaling, and may discourage the uncritical use of global reanalyses for applications requiring high spatial detail. If this is part of the authors' intent, I would encourage them to state it explicitly.

We thank the reviewer for this valuable comment. We agree that remapping all products to a common coarse-resolution grid would be an appropriate strategy if the sole objective of the intercomparison were to assess consistency among datasets.

However, as discussed in our response to Comment 5, the evaluation framework adopted in this study has a broader objective. In addition to assessing the consistency of MORE relative to observations and existing reanalysis products, we also aim to evaluate the dataset at the spatial scale for which it was specifically developed and intended to be used. In this context, part of the motivation is indeed to investigate the extent to which kilometre-scale convection-permitting downscaling contributes to the representation of fine-scale atmospheric variability and spatial structures that may not be fully captured by coarser-resolution products.

We agree that this rationale was not sufficiently explicit in the original manuscript. Following the reviewer's suggestion, we will clarify the methodological motivation behind the adopted validation strategy and the interpretation of the intercomparison results, explicitly distinguishing the consistency assessment aspect from the objective of evaluating the potential benefits of convection-permitting downscaling.

Comment figure 13 :

Figure 13. The use of dots renders the figure somewhat difficult to interpret. The authors may wish to consider replacing dots with lines, which would sacrifice some precision in the representation of individual values but would likely improve the readability of the distributional tails considerably

We thank the reviewer for this suggestion. We agree that readability of the distribution tails is important and we will further explore possible improvements to the graphical representation. While preserving the information conveyed by the current PDF representation, we will investigate whether alternative plotting choices can improve the readability of the figure in the revised manuscript.

Comment fiugure 14 :

Figure 14. Change MOLOCH-ISAC into MORE.

We thank the reviewer for noticing this inconsistency. The label "MOLOCH-ISAC" in Figure 14 will be replaced with "MORE" in the revised manuscript.