

Response to Minor Revisions – "Cell tracking -based framework for assessing nowcasting model skill in reproducing growth and decay of convective rainfall"

Jenna Ritvanen, Seppo Pulkkinen, Dmitri Moisseev, and Daniele Nerini

January 16, 2025

We thank the reviewer for their constructive comments on our submitted manuscript. We have copied the comments from the reviewer and the topic editor in black here and include our response to each individual comment in blue. We also include a `latexdiff` version of the manuscript to this response that shows the detailed changes from the previous submission.

Reviewer 1

The paper has made modifications in response to the previous review comments, primarily adding comparisons of the proposed framework with pixel-by-pixel evaluation metrics, highlighting its advantages and limitations. It also explains the differences between the proposed framework and traditional pixel-level validation methods and offers a multi-dimensional analysis of the performance of different precipitation models, providing new insights for precipitation forecasting model evaluation. However, the response to the review comments is still incomplete, and some areas of the revision need further improvement. The following points should be addressed:

1. For the last review mentioned that the article lacks enough information to evaluate the life cycle of cells. The authors should improve more evaluation metrics and visual methods such as tables.

Thank you for the comment regarding evaluating cell life cycles. We note that the aim of the framework is not to study the life cycle of convective cells, but rather evaluate how well the nowcasting models reproduce the development of the cells after the nowcast is created. For this purpose, we believe that the metrics presented here capture most of the cells in the dataset sufficiently. Even though the forecast horizon is limited to one hour, according to Figure 4j, a majority of the cells in the dataset have a lifetime of less than 1 hour after the nowcast is created. Additionally, none of the models evaluated in this study are expected to predict realistic cell development for longer time periods, since the models (apart from the advection nowcast, which predicts no development) produce increasingly smoothed nowcasts as the leadtime increases.

However, in future studies it might be necessary to adjust the framework increasing the forecast horizon to capture the full lifecycle of all cells after the nowcast is created, e.g., to evaluate performance of models that are expected to predict convective cell development for longer time periods and/or do not

produce smoothing, or if majority of the dataset is large, longer-living cells or convective systems. This is mentioned as possible future research in the discussion.

2. It is necessary to revise the framework description and address its limitations.

Thank you for the comment. We have improved the discussion of the limitations and uncertainties of the framework in Section 5.

3. It is beneficial to provide more in-depth comparative experiments and discussions after given the advantages and limitations observed after comparing with pixel-level metrics.

We have added discussion comparing the pixel-level CSI to the cell-based CSI in Section 4.3.1. As the manuscript is already quite long and the purpose is to provide an initial description of the framework, we feel that any additional experiments comparing the framework with pixel-level metrics would increase the manuscript length overtly without providing much benefit. As such, we leave further experiments for future studies.

4. In chapter 4.2, why CSI and RMSE were selected for the pixel-level evaluation metrics instead of alternatives like MAE or MSE?

The CSI and RMSE metrics were selected in order to provide metrics for direct comparison with the metrics calculated from the cell tracking -based framework.

Pixel-based CSI provides an opportunity for comparison with the different CSI versions used in the framework (e.g. when evaluating how well the nowcasts reproduce cell existence). We have improved the discussion of the CSI differences in Section 4.3.1.

RMSE was selected over similar metrics such as MAE or MSE because, as an L2-norm based error, it penalizes high error values more than MAE (similarly to MSE), but compared to MSE provides the error value in original units. We also calculated results for other possible error metrics and found no substantial differences in the results between the models, and therefore concluded that including RMSE provides sufficient information for the purposes of this study. However, in future studies, the metrics should of course be selected to be suitable for the goals of each study.

Editor comments

List the various sources of uncertainties that could impact your results. Also suggest how the uncertainties could be addressed or considered in a future related research.

We have added a paragraph in Section 5 discussing the uncertainties that can impact the results.