**Note to Dr Yuefei Zeng**,

While we were preparing the response to this comment (a first draft of which you will find below), we received an email from Mr. Bastien Chatelon, who made us aware of a significant bug in the data assimilation algorithm published on Github and linked to this paper. All the simulations reported and analysed in the current version of the manuscript are affected by this bug.

Although we believe that the bug does not alter the value of the research conducted, nor the substance and content of our work, we do think that the best way forward at this point is to rerun the tuning of the system from the start and present a new configuration (which will likely entail a new set of optimal parameters and therefore new figures) once the issue has been fixed. This process may take some time, especially as the authors are currently busy in other tasks related to their current occupations and jobs.

As a result, we have decided to withdraw the current version of the manuscript from the peer-review process and to submit a revised version at a later date.

The authors.

## Response to 'comment on gmd-2022-269' by Dr Yuefei Zeng

We would like to thank Dr Yuefei Zeng for his interesting and relevant comments. We have addressed his remarks below, on a comment-by-comment basis. We have also indicated in the text how we intend to modify or integrate the manuscript accordingly.

**Editor, comment #1:** Since this study is very close to my work, I have read it with great interest and would like to provide some comments for your consideration.

1. I wonder why rotation is omitted in this work. The model can not run with both rotation and orography (it is found in the code that there is no initialization routine for such case)? Since authors attempt to assess the relevance of this study for the real-data convective-scale data assimilation, it would be very interesting and even necessary to show the impacts on rotational and divergent part of dynamics (A similar study but for large-scale data assimilation study has been done by Zeng and Janjic 2016, showing the importance of reconstruction of the rotational part). However, the behavior may be different from the large scale. From this point of view, a study with rotation may be more appealing than with orography (Note: there is no discussion on importance of orography in this study).

**Response to comment #1:** As Dr Zeng has rightly noticed, rotation is not considered in the experiments described in this paper, although it is indeed included in the modRSW model, and has been used both in the case studies considered in the initial model description paper already published (Kent et al., 2017), as well as in the satellite data assimilation experiments conducted by one of the authors in his PhD thesis (Cantarello, 2021, available at this link). In order to consider rotation and orography together, substantial modifications to the numerical scheme currently used would be required. For this reason it was not possible to include

rotation in a straightforward way in the case study illustrated in this manuscript. The possibility of conducting simulations with both rotation and orography would certainly constitute an interesting model upgrade, although we believe it lies outside the scope of this work. We will revise the introduction in the manuscript to clarify more explicitly this limitation and provide a better justification for our choice.
Regarding the orography, we note that it plays a very important role in obtaining continuous generation of gravity waves which in turn trigger convection and precipitation in our periodic domain without having to reinitialise the experiments. We will clarify the text in section 4.1 of the manuscript to highlight this point better.

**E, c2:** 2. Experimental settings: 1) The observational resolution is 50 km, which is much coarser than that of observations usually used in convective-scale data assimilation, e.g., resolution of Doppler radars ~ O(1 km). Therefore, from this point of view, it can not be considered as convective-scale data assimilation.

**Response to comment #2:** Dr Zeng correctly points out that the common observation density in convective-scale data assimilation systems - when they include radar observations - is higher than 50 km. However, in our experiments we have chosen to simulate only conventional surface observations, for which a 50km station spacing is fairly realistic.
Nonetheless, to address this comment and a similar one made by reviewer 1, we have decided to modify the observing system, including the assimilation of some pseudo Doppler radial wind observations.

**E, c3**: 2) On one hand, to evaluate the skill of spread, a metric called ”the spread skill ratio” (Aksoy et al., 2009) is often used, which is calculated as (Spread+Obs. Error)/RMSE. It is optimal if equal to 1. In this study, Spread/RMSE is used, which neglects the observation error and therefore results are probably underestimated. However, on the other hand, considering that the (truncation) model error could be accounted for by the specific additive inflation (the similar method is applied in Zeng et al. 2019, Zeng 2020), additional use of inflation like relaxation methods may result in overestimated spread, which may compensate the deficiency of Spread/RMSE. In my opinion, (Spread+Obs. Error)/RMSE should be used, and the tuning interval of relaxation factor should be efficiently chosen, currently a great part of experiments with larger relaxation factors are not necessary. To my experience, the additive inflation may be already sufficient to maintain the spread (A related study can be found in Zeng et al. 2018).
**Response to comment #3**: We thank Dr Zeng for his remark about the spread skill ratio. Aksoy et al. (2009) measure spread in observation space and compare this with the RMS of innovations based on the ensemble mean in observation space. The latter quantity involves noisy observations and therefore requires the observation error to be taken into account. We agree that observation error needs to be accounted for when observations are included in a performance metric. For example, Bowler et al. (2017) [https://rmets.onlinelibrary.wiley.com/doi/full/10.1002/qj.3004, Appendix B] adjusted the RMSE measure to take account of observation error when verifying against observations. However, in this study, we compare the spread in model space with the RMSE of the ensemble mean, calculated using the nature run as a truth reference. No observations are involved in this metric, so the inclusion of observation error is not required in this instance.
Regarding the use of combined additive and multiplicative inflation, we note that by including both additive inflation and RTPS relaxation in the tuning experiments summarised in Figure

5, we are implicitly showing that optimal performance requires both elements for this experimental configuration. Nonetheless, we plan to add a citation to both Zeng et al. papers (2018, 2019) in section 3.3.2, noting that Zeng et al. (2018) differs in using Q = B (climatological background-error covariances), which combines forecast and observation errors.

**E, c4:** 3. Miscellaneous: 1) There are adaptive RTPP and RTPS but the ones used in this study are obviously not adaptive (see Kotsuki et al. 2017; Ying and Zhang, 2015). 2) How is additive inflation Gaussian? and why are correlations are ignored? We have done similar study (Zeng et al. 2019), those additive perturbations are not Gaussian and we have not explicitly removed the correla- tions. 3) The typical lead time length for convective-scale data assimilation is 6 hours, instead of 3 hours used in this study. Furthermore, to validate forecasts, it would be more informative to see the plots of variations of RMSE with the lead time.

**Response to comment #4:** The aim of this study is to demonstrate that the modRSW model is a credible tool for convective-scale investigations, rather than to produce an optimal data assimilation system. We agree that the additive inflation with a diagonal model-error covariance matrix Q is not optimal, and believe that investigations to find better ways of implementing additive inflation constitute valuable research. Indeed, we hope that the modRSW model and code might provide a framework for such future investigations.
Regarding the definition of nowcasting, we note that the World Meteorological Organization (WMO) Working Group on Nowcasting Research has defined nowcasting as being a period from the present to six hours ahead. Indeed, the document at this link refers to an example of nowcasting in the 0-2 hour range.
To address the final point, we will include a plot of the variation of RMSE with lead time.

**E, c5:** To sum up, I believe that this framework can be very useful for the community of convective-scale data assimilation, but the presentation needs to be greatly modified.

**Response to comment #5:** We thank Dr Zeng for his constructive suggestions.