The reviewer's comments are presented in black, while our answers are in blue.

Response to Reviewer 2

General comments:

This study aims to demonstrate that the assimilation of Aeolus in the model used as boundary conditions for the dust simulations improves the capability of the regional model to resolve dust loads.

The paper is readable and the results are validated from external measurements.

The study is based on several months of simulation but only focuses on a reduced period and one particular event. Studies with Aeolus data on other specific events (e.g., tropical cyclones) acknowledge that it is difficult to work on such a reduced set of events (see, for instance, DOI: 10.1002/qj.4370, The characterization and impact of Aeolus wind profile observations in NOAA's regional tropical cyclone model (HWRF) by Marinescu et al., 2022). The impact is not systematically in the direction of the average impact (i.e., assimilating Aeolus data can also be very detrimental in some cases).

I think this paper requires such a discussion on the significance of the results when considering the reduced set of events.

Without this discussion, this paper is still an important contribution to the demonstration of the usefulness of Aeolus data, through a well-documented case study. However, I don't think it stands as a solid proof by itself.

Reply -> We would first like to thank the reviewer for their time and constructive feedback. The paper does not aim to solely and uniquely evaluate the potential of assimilation throughout the lifetime of Aeolus but rather stands as a test case and proof-of-concept, and this is noted throughout lines 480-484 where various published papers and ongoing studies evaluating the performance of Aeolus data are mentioned. To correctly depict the significance of the paper the following sentence has been changed from "Concluding, the incorporation of the Aeolus products improves the predictive ability of the WRF-Chem model for the East Mediterranean and Middle East regions, by reducing positive bias and underestimates." to "The benefits attained from the incorporation of Aelous, solely regard the period of 14 to 25 October 2020, where anticyclonic conditions prevail in the EMME and Central Mediterranean regions. Even though the period of improvement is statistically negligible compared to longer timescales, the strong reductions in positive bias and underestimates highlight the importance of Aeolus in further dust research.".

Specific comments:

- I do not understand the expression "comparisons with the whole simulated domain diffused the improvements" (l. 355) or "Statistical comparison of all 56 AERONET stations within the extended model domain diffuses the improvement" (l.436) and another occurrence on l. 468. In particular, this use of the word "diffuse". This might

be a specific jargon that I am not aware of, but could you explain this notion with a different wording?

Corrected -> Removed the use of the word diffuse throughout the manuscript

- l.88 Is “turmoiled” necessary?

  Reply -> Turmoiled was placed to demonstrate the effect political instability has in the world of academia, but has now been removed as requested.

- l.109: Did you mean “in section 2.1 to 2.5”?

  Reply -> Corrected to clearly reflect what is described per section.

- l.197 “miscrophysical”

  Corrected

- l.260 “deterrent”, not sure I understand, did you mean “inherent”?

  Corrected

- l.263: why was the nearest hour so different from a 3-hour average? Is the model AOD noisy or extremely variable?

  Reply -> In short, the anticyclogenesis present during the study period would see a highly variable AOD. There are 3 to 5 MIDAS overpasses with a 5-minute gap between them. The closest hour approach is configured to compare MIDAS values that fall within hh:00 to hh:29 minutes to the WRF interval hh, where hh refers to the hour. For values at hh:30 to hh:59 the WRF interval of hh+1 is used in the comparison. Hence, if MIDAS had an overpass during 13:25, 13:30 and 13:35, the nearest hour approach will solely compare the 13:25 interval to the WRF product at 13:00 and the other two to the WRF product at 14:00. Thus, the rolling 3-hour approach would be more adept to capture and compare this variability relative to the nearest hour approach. Additionally, the 3-hour approach is more aligned with the weighted average approach, which uses two model hours and weights them according to whether the overpass was after hh:30 or before hh:30 (Eq. 3).

- l.265-269: that’s a lot of averaging. How different are the values where there is an overlap for instance? In addition, I don’t understand what is being produced here. Maps of MODIS AOD?

  Reply -> There are a total of 2,192, 5-minute, MIDAS retrievals that do not continuously cover the whole domain. To ensure all the available observations were used, daily sums were created with overlaps averaged. Then the daily sums were averaged to produce a single AOD map, being the most effective way of qualitatively visualising the comparison. Stated on lines 272 and 274 AOD maps comparing the collocated products were produced.

- l.295: Please describe better figure A2: It shows that both model runs give very similar results at the Agia Marina station during the spring period.

  Correction -> Clarified what figure A2 points at.

- l.305-306: statistical significance again… This contradicts the l.308 statement of a "thorough investigation"

  Reply -> Following lines 303-308 the statistical significance refers to the comparison with EMEP. While line 310 foreshadows the thorough investigation to follow for the stated period. Changes have been made to remove any contradiction.

- Fig. 5: I would suggest to either remove hel1 and hel4 from the legend or introduce it somewhere in the text.

  Corrected

- l.334: Isn't it 4 FLEXPART runs but only "two, 5-day periods"?

  Corrected

- l.342: Fig A4 does not show the AERONET stations

  Corrected

- l.411: LIVAS is a dataset, not a "lidar"

  Corrected

- Fig 8: Is it possible that dust events happening close to the domain boundary are less well resolved? (e.g. dust could be transported from outside the domain, across the boundary). There are also some discrepancies to the East of the Caspian Sea for instance. The other hypothesis would be that getting the magnitude wrong on strong events already produces a large error. And unfortunately, they happen close to the domain border for this period.

  Reply -> Dust events close to the domain boundaries are less resolved, in particular, discrepancies East of the Caspian Sea are noted. The relative error of simulating events increases with the magnitude of the event, hence strong events close to the boundaries inherit a larger error relative to events in the inner domain, which is the case for this study period. This is now included in the text. Additionally, referring to Fig. 8c, location 1 the Bodele depression, has been the subject of various past model simulations identifying an inability of various models to accurately depict dust mobilisation in the locality, also included in the text.