

The article has improved from the previous version. I broadly accept the authors' responses to my previous comments.

With that said, there are still numerous locations in the manuscript where the text needs to be modified either for clarity or to better describe the results obtained. I examined the tracked-changes version of the manuscript and found some areas where the text was changed that I'm not sure needed to be. I presume those changes were the result of other reviewer comments, which I will not contest.

Please make the following changes to the text. If you make these changes appropriately, there should be no need for me to review an additional iteration; that is, the subsequent version should be acceptable as-is for publication:

- (1) Line 122: should the units of total number concentration not be m^{-3} instead of kg^{-1} ?

Reply: revised as suggested.

- (2) Eqn (6)/Line 156: Since this is merely a constant and the parameters are not obvious nor explained, you may open yourself to additional complaints by including these intermediate terms (or not defining them). I think in this instance it is acceptable to eliminate the intermediate formulation and simply state the value of the constant.

Reply: The equation has been revised to avoid introducing additional undefined parameters while preserving the formulation used for the dry-snow reflectivity calculation.

- (3) Lines 172-182: I suggest you incorporate a portion of your response to my previous comment #6 on explaining which test case these data were derived from. I accept that the particular case may not matter, but it would still be nice to inform the readers that (1) the data obtained here are not from the main case in the study, and (2) the specific case shouldn't matter anyway since these results should be independent of weather regime. Even if this is obvious to the authors, it may not be so to the readers.

Reply: We have incorporated the relevant part of our previous response into the manuscript. Specifically, we now state that the fitting dataset was derived from a 2-h WRF forecast snapshot valid at 2000 UTC 20 May 2019 and is independent of the 13 May 2022 derecho case examined in the main experiments. We also clarified that the choice of fitting case should not materially affect the resulting operator because the fitted relationships represent the mapping between hydrometeor fields and reflectivity implied by the

Thompson microphysics formulation, rather than the dynamics of a specific weather regime.

(4) Line 175: "a widespread ...(missing noun)" is poor language, as "widespread" is an adjective, whereas "wide spread" is a noun phrase. Change to "large spread" to avoid this ambiguity.

Reply: revised as suggested

(5) Lines 179-181: repeated verbiage. Delete "introduces separate...dry and wet snow", change "resulting" to "results," and delete "and substantially improves the agreement". Also, on line 181, which section of the plotted data points in Fig. 1a represents wet snow? Either annotate this portion, describe it in the text, or just eliminate the statement, "particularly for the wet-snow branch."

Reply: revised as suggested

(6) Line 182: change "considerable" to "some"

Reply: changed

(7) Lines 293-294 and Figure 4: Based on this statement, the innovation, then, must be formulated as $\text{innovation} = y - H(x)$, which is standard. But **all** of the experiments show positive innovations in Fig. 4. This statement doesn't draw a distinction between NODA and the other experiments. Looking at the tracked-changes version of this manuscript, I prefer the old statement to this one. Did another reviewer ask you to change it? If so, then you can keep the present version, but you still need to draw a distinction to the other experiments.

Reply: We have revised the text to state explicitly that the innovations are defined as observation minus model equivalent, that NODA has the largest positive innovations, and that the radar-assimilation experiments substantially reduce, but do not completely eliminate, the positive mean innovations during the cycling period.

(8) Lines 303-305: This is not the correct way to state the comparison. JEDI-org is quite a bit higher than JEDI-New and GSI-new in Fig. 4d. In Fig. 4h the three experiments are essentially identical. Change the wording to reflect this.

Reply: We agree with the reviewer. The previous wording incorrectly combined the RMSI and mean-innovation comparisons. We have revised the text to state that JEDI-Org has noticeably larger analyzed-(V_r) RMSIs than JEDI-New and GSI-New (Fig. 4d), whereas the analyzed-(V_r) mean innovations of the three radar-assimilation experiments are nearly identical and remain close to zero (Fig. 4h).

(9) Line 310 as well as 388-392: I don't know what the underlying theoretical/mathematical differences are between how GSI computes the analysis and how JEDI does it, but if the two methods are similar, that would also explain the similarity. This is worth noting in the text.

Reply: We have revised the text to clarify that the comparable behavior of JEDI-New and GSI-New may be attributable not only to the use of the same updated reflectivity observation operator, but also to the broadly similar ensemble-based variational formulations used in the two systems. In particular, both systems use flow-dependent ensemble covariances to spread observational information to the model state variables, which can help explain the similarity in the resulting hydrometeor increments and forecast performance.

(10) Lines 386-387: replace "mid- to upper-levels" with a specific range of geometric heights. Do not mention the native model vertical levels, as readers will not be able to link those to physical heights.

Reply: changed to 4-12 km

(11) Line 388: insert a space in "...GSI-New(which uses..."

Reply: Revised

(12) Lines 400-401: Thank you for acknowledging the non-negligible qr increments that I brought up previously. I request that you add/clarify some details. First, I think it is worth noting that the positive qr increments are restricted to below the 2-3-km level, i.e., below

the freezing level. In addition, the q_s and q_g increments are negligible in this same region, so indeed it is the rainwater increments that rebuild the missing western tower, a success! Additionally, I think it is also worth noting that there are non-negligible negative q_r increments in the eastern tower and presumably above the freezing level (looks to be about 3-5 km) in JEDI-New. Even though the negative q_g increments at this location clearly dominate, it is still the case that JEDI-New was able to adjust q_r to correct for the overly strong reflectivity in the background, which is not something that is seen for the other experiments. This is another way for you to show off how great of an improvement the new operator and JEDI are for this process. Take advantage of this!

Reply: We thank the reviewer for this helpful interpretation. We have revised the text to note that the positive rainwater increments in the western tower are confined mainly below about 2–3 km, below the approximate freezing level, where snow and graupel increments are negligible, indicating that the rainwater increments help rebuild the missing western tower. We also added discussion of the non-negligible negative rainwater increments in the eastern tower of JEDI-New at approximately 3–5 km, which help correct overly strong background reflectivity together with the dominant negative graupel increments.

(13) Line 423: "unrealistically" opens the door for philosophical debate. I suggest deleting the rest of this sentence after "Minnesota."

Reply: deleted

(14) Lines 423-425: in the sentence following that from comment (13), I suggest a different way of expressing the result. Try this: "NODA also generally overdoes coverage of convection across Minnesota and Iowa in forecast hours 3 and 5."

Reply: We have revised the following sentence to state that NODA generally overpredicts the coverage of convection across Minnesota and Iowa at forecast hours 3 and 5.

(15) Lines 466 & 467: there is an extra space after "radar DA" (466). On 467 change "other more directly impacted fields" to "the directly modified fields"

Reply: Revised as suggested

(16) Line 476: I think it is helpful to note the specific locations where this is valid [i.e., insert "(especially northeastern South Dakota and southwestern Minnesota)"]

Reply: We have revised the sentence to specify that the stronger surface winds behind the gust front in the DA experiments are most evident over northeastern South Dakota and southwestern Minnesota.

(17) Line 496: delete "tested" and the parentheses

Reply: Revised

(18) Line 504: change "45 dBZ case" to "45 dBZ threshold"

Reply: Revised

(19) Line 651: I don't see a reference to the appendix anywhere within the manuscript body. Check whether such a reference is required and insert it somewhere if so.

Reply: We thank the reviewer for pointing this out. We have checked the manuscript and inserted a reference to the appendix in section 3.2.2, where the En3DVar data assimilation system is introduced. The text now states that a brief formulation of the incremental En3DVar used in this study is provided in the appendix.