Response to Reviewer 3

The authors thank Anonymous Reviewer #3 for his or her time in reviewing our manuscript, and for their helpful comments. Addressing them has enabled us to improve this manuscript. Our point-by-point responses are indicated in blue below.

General comments

This manuscript uses the HAFS model with the GSI 4DEnVar system to examine the impacts of assimilating ROMEX, COSMIC-2, and lower-level radio occultation observations on hurricane forecasts. The ROMEX dataset includes multiple RO sources, with the commercial observations coming primarily from Spire and PlanetiQ. The study performs cycling data assimilation and forecasting for four hurricane cases and further presents a more detailed case analysis of Hurricane Fiona (2022). Overall, the topic is valuable for assessing the potential benefits of commercial RO data for hurricane forecasting. However, several aspects of the interpretation and mechanistic discussion would benefit from further clarification to strengthen the rigor and overall persuasiveness of the study.

Specific comments

1. Figure 2 shows the GNSS RO observations available for assimilation for Hurricane Ian, whereas the more detailed case analysis in Section 5 focuses on Hurricane Fiona. The rationale for choosing Ian rather than Fiona in Fig. 2 is not entirely clear. If Fiona is the primary case for the subsequent detailed discussion, it may also be useful to provide the corresponding observation-availability information for Fiona, which would improve the continuity of the manuscript and better support the later analysis.

We appreciate this feedback, and we agree that it's better to show profile counts for Fiona to improve continuity of the manuscript and interpretation of results shown in Section 5. In an earlier draft of this figure, we used a 2 x 2 multi-panel plot showing the profile-per-cycle counts for all four cases. However, fitting all four panels onto a typeset page required shrinking the panels to the point where the individual bars became difficult to distinguish. Therefore, we decided to show time series of profile counts for just one case. We had previously selected Ian to show because it was located deeper in the tropics during its earlier period, enabling us to show more similar per-cycle COSMIC-2 and Spire counts. But it is well known that COSMIC-2's sampling density peaks near the Equator, and we agree that it is better to show Fiona for the reasons you mention. Because Fiona has more DA cycles than Ian, we decided to compare the "available profile" and "pass QC below 800 hPa" profile counts in two separate panels, (a) and (b) respectively, to improve visual clarity. The re-plotted Fiona Figure 2, shown below, will replace the old Ian Figure 2 our revised manuscript.

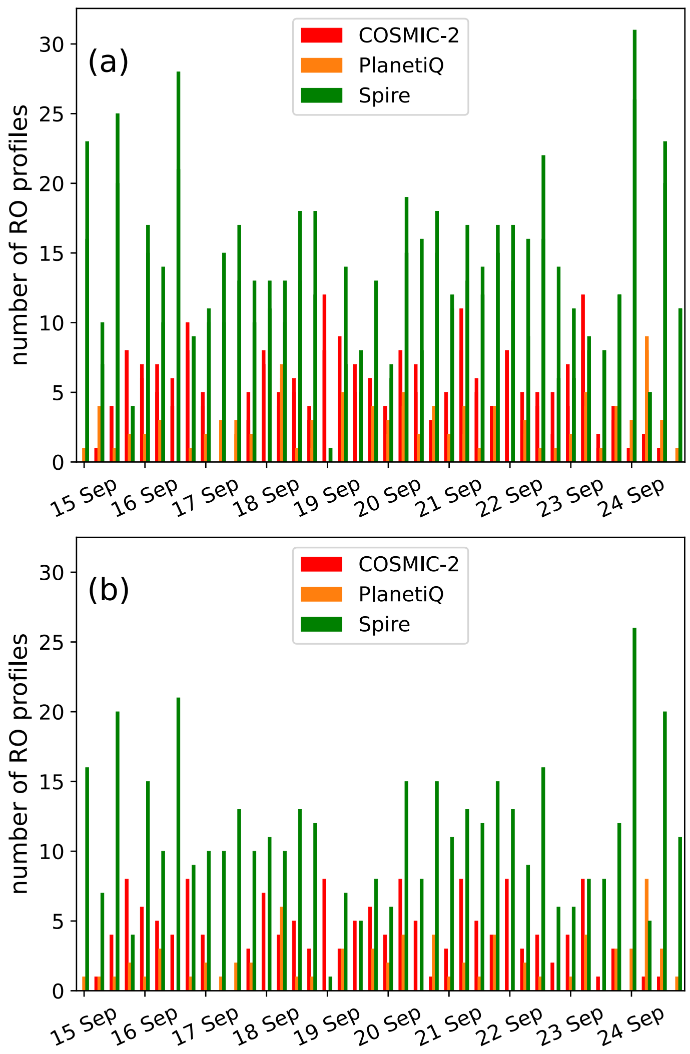

Figure 2. (a) Number of COSMIC-2, PlanetiQ, and Spire GNSS RO profiles available for assimilation in the HAFS-A moving nest for individual Hurricane Fiona (2022) DA cycles, shown as red, orange, and green bars, respectively. Each 0000 UTC cycle time is labeled with its date on the *x*-axis. (b) As in (a), except showing the number of profiles with at least one bending angle assimilated below 800 hPa.

Additionally, following Reviewer 2's suggestion we added a new Table 1 to the manuscript that lists the number of assimilated profiles per day for each TC case.

2. In Fig. 3c, PlanetiQ, like Spire, is also characterized by lower SNR than COSMIC-2. It is therefore somewhat unclear why only Spire exhibits a markedly lower rejection rate, whereas PlanetiQ does not show a similar behavior. The authors are encouraged to elaborate on the possible reasons for this difference.

One possible reason for the differences in extreme O-B outlier frequency between PlanetiQ and Spire, which would affect their OMB rejection rates relative to COSMIC-2 shown in Figures 3c and 3d of the previous manuscript (now panel 3b in the re-plotted Figure 3), could be differences in the methodologies applied by PlanetiQ and Spire when processing or running quality control checks on their data before sending it to EUMETSAT.

As we explain in our AC response to Reviewer 2's Major Comment 4, we decided to remove discussion in this paper relating the PlanetiQ, Spire and COSMIC-2 diagnosed Statistical Check rejections to their known SNR characteristics. One the one hand, COSMIC-2 SNR has a latitudinal dependence due to its satellites' orbital paths and antenna geometry (Ho et al. 2023 Fig. 5). Additionally, GNSS RO BUFR data files do not contain SNR information, and so we cannot evaluate the SNR statistics for the subset of global data assimilated in our experiments. We also are aware of the fact that commercial RO data providers may not be willing to sell their highest-quality data with the highest SNR to government customers at a mutually agreed-upon price.

We also re-plotted Figure 3 to move the rejection percentages of all three missions onto the same panels to facilitate comparison.

Reference:

Ho, S.-P., X. Zhou, X. Shao, Y. Chen, X. Jing, and W. Miller, 2023: Using the commercial GNSS RO Spire data in the neutral atmosphere for climate and weather prediction studies. *Remote Sensing,* **15,** 4836.

3. Observation-error estimation generally requires sample statistics accumulated over a sufficiently long period. However, Fig. 4 does not clearly indicate whether these results are derived from a single hurricane case or from statistics aggregated over all four cases. Even in the latter case, the sample still appears rather limited for the results to be interpreted as robust and broadly applicable observation-error estimates. I therefore encourage the authors to clarify that these are local sample statistics or local diagnostic estimates to avoid possible misunderstanding by readers.

Our diagnosed observation-error statistics shown in Fig. 4 are drawn from the full set of ROMEX experiment analyses taken from all four TC cases. Thank you for pointing out this need for clarification. Yes, we agree that caution is warranted when extrapolating these statistics from a relatively small local sample to more general statements comparing the uncertainty characteristics of the different RO datasets. We revised this paragraph and it now reads, with revised portions in bold, as (lines 409-420):

"Figure 4 compares profiles of COSMIC-2, Spire, and PlanetiQ observation error standard deviations estimated from the ROMEX experiment's O-B and observation-minus-analysis (O-A) **data collected from all four TC cases**. **These diagnosed observation errors** include contributions from instrument, retrieval, and representative sources, **and we estimated them using the Desroziers et al. (2005) algorithm, which has been widely employed by operational NWP centers (Cucurull et al. 2013; Bonavita 2014; Bowler 2020; Lien et al. 2021). Some caution is warranted in drawing broader conclusions about the uncertainty characteristics of these three RO observation platforms, given our relatively small DA diagnostic data sample drawn from a TC-following regional domain. From our experiments,** we find that all three platforms have similar observation uncertainty in the upper troposphere. However, Spire observations show the smallest uncertainty below the 5-km impact height in the tropics (Fig. 4a) and below the 7-km impact height in the extratropics (Fig. 4b)."

4. Figures 5 and 6 contain several statistically significant results that do not appear fully consistent with the broader implication of the manuscript that assimilating more observations is generally beneficial. For example, while it is understandable that Control_noC2 exhibits larger errors than Control in Fig. 5, Control_noC2 also shows statistically significant relative skill improvement at 60 and 72 h in Fig. 6. In addition, Fig. 7 suggests that the Pmin bias in Control_noC2 is overall closer to zero than that in Control. Taken together, these results appear to suggest that excluding COSMIC-2 may, in some respects, lead to better forecast performance, which seems somewhat at odds with the broader interpretation presented in the manuscript. Further clarification from the authors would therefore be helpful.

You raise an important point here. In our opinion, despite some negative impacts, such as the small medium-range forecast degradation in Pmin bias and absolute error, our results still show that assimilation of COSMIC-2 on top of the other government missions is overall beneficial in HAFS for this set of four TCs. In the Summary and Conclusions, we highlight one possible reason why we find examples of small or mixed impacts with COSMIC-2 assimilation in this study (lines 782-789), where new words inserted with this revision are highlighted in bold:

"Also, we should note that the statistical forecast error analysis described in Section 4 was heavily influenced by the long-lived Earl and Fiona cases, which spent most of their time as hurricanes outside the tropics. Considering the marked decrease in COSMIC-2 profile density with latitude in the extratropics (Ho et al. 2020b), the incremental benefits of assimilating additional RO data may be proportionally greater **for TCs moving** outside the tropics, given the more uniform latitudinal distribution of sun-synchronous polar orbiting Spire and PlanetiQ satellites (Anthes et al. 2024). **Future studies evaluating larger TC samples could test this hypothesis.**"

The decreasing COSMIC-2 sampling density with latitude outside the tropics is reflected in our manuscript's new Table 1 (copied below), which shows the number of RO profiles assimilated in the HAFS nest per day for each experiment. The increase in RO profiles from Control_noC2 to Control becomes markedly smaller for the Earl analyses on 10 Sep, Fiona analyses on 24 Sep, and Ian analyses on 01 Oct, for which the domain was centered poleward of 30 degrees N. As part of our response to a related comment by Reviewer 2, we re-ran our forecast track and Pmin mean absolute error calculations for the subset of 42 analyses initialized equatorward of 42 degrees N (Earl through 18 UTC 06 Sep, Fiona through 12 UTC 21 September, Ian through 18 UTC 27 September, and all Julia analyses). For this analysis subset, removal of COSMIC-2 from the Control configuration leads to greater degradation of TC track forecasts, as Control_noC2 now shows about negative 5-10% relative skill between the 54 and 96 h lead times (Fig R3.1), whereas for the full analysis sample shown in Figure 5 the Control_noC2 track forecast relative skill stays within ±5% for most lead times.

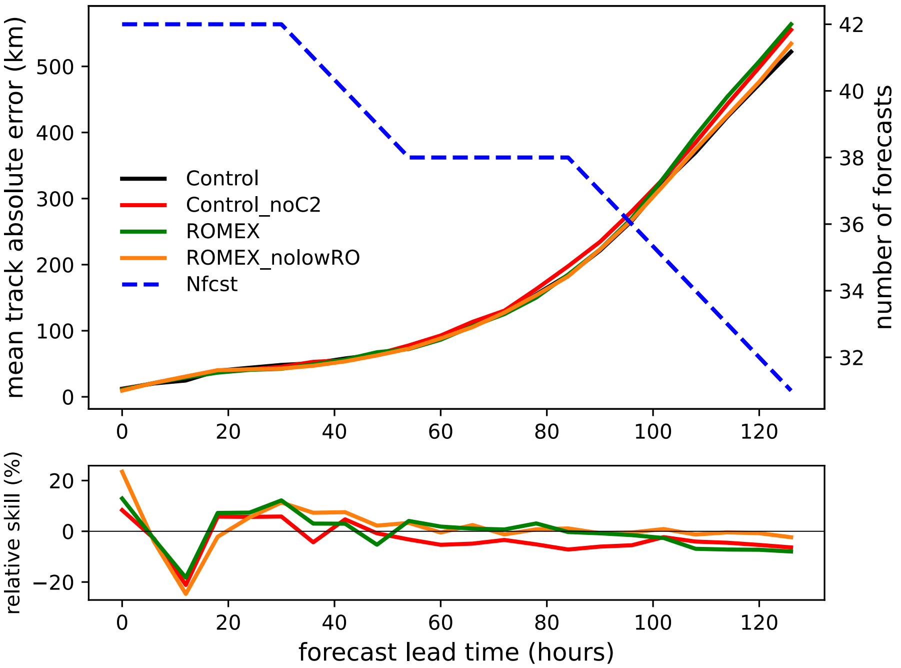

Figure R3.1 As in the manuscript's Figure 5 but limiting the TC position error sample to forecasts initializing the TC center south of 24 degrees N. Note the larger *y*-axis scale for relative skill shown in the lower panel, compared to Figure 5.

We added some further emphasis of the positive COSMIC-2 impacts found in this study to the manuscript. In Section 4a we added this new sentence, highlighted in bold, mentioning the results shown in Figure R3.1 (lines 448-452):

"Control_noC2's negative relative skill during the same period (Fig. 5b) indicates that COSMIC-2 observations are also beneficial for TC track forecasts. **This is particularly true for the subset of analyses that initialize TCs equatorward of 24° N, for which removing COSMIC-2 from Control yields negative 5-10% relative skill during *t* = 54-96 h (not shown)**."

Also, in the Summary and Conclusions section we added this sentence (lines 789-793):

"Additionally, by showing Control_noC2's short-range intensity and medium-range track forecast degradations relative to Control, this study supports previous work showing positive impacts of COSMIC-2 assimilation on regional model TC forecasts (Miller et al. 2023; Teng et al. 2023) and the need to sustain COSMIC-2-like coverage over the tropics in future RO satellite missions."

N_{cyc}	**Earl**			**Fiona**			**Ian**			**Julia**		
Date (n_{cyc})	noC2	Control	ROMEX	noC2	Control	ROMEX	noC2	Control	ROMEX	noC2	Control	ROMEX
03 Sep (3)	1	19	73	—	—	—	—	—	—	—	—	—
04 Sep (4)	2	24	92	—	—	—	—	—	—	—	—	—
05 Sep (4)	4	22	83	—	—	—	—	—	—	—	—	—
06 Sep (4)	4	22	86	—	—	—	—	—	—	—	—	—
07 Sep (4)	4	17	69	—	—	—	—	—	—	—	—	—
08 Sep (4)	3	17	101	—	—	—	—	—	—	—	—	—
09 Sep (4)	3	27	97	—	—	—	—	—	—	—	—	—
10 Sep (4)	4	13	101	—	—	—	—	—	—	—	—	—
15 Sep (4)	—	—	—	4	17	87	—	—	—	—	—	—
16 Sep (4)	—	—	—	6	30	104	—	—	—	—	—	—
17 Sep (4)	—	—	—	4	14	80	—	—	—	—	—	—
18 Sep (4)	—	—	—	1	24	97	—	—	—	—	—	—
19 Sep (4)	—	—	—	6	40	85	—	—	—	—	—	—
20 Sep (4)	—	—	—	5	27	101	—	—	—	—	—	—
21 Sep (4)	—	—	—	3	29	100	—	—	—	—	—	—
22 Sep (4)	—	—	—	5	28	102	—	—	—	—	—	—
23 Sep (4,3)	—	—	—	5	30	82	4	29	58	—	—	—
24 Sep (4,4)	—	—	—	6	10	96	9	42	94	—	—	—
25 Sep (4)	—	—	—	—	—	—	4	31	79	—	—	—
26 Sep (4)	—	—	—	—	—	—	4	19	58	—	—	—
27 Sep (4)	—	—	—	—	—	—	6	18	84	—	—	—
28 Sep (4)	—	—	—	—	—	—	5	25	105	—	—	—
29 Sep (4)	—	—	—	—	—	—	4	21	103	—	—	—
30 Sep (4)	—	—	—	—	—	—	5	25	84	—	—	—
01 Oct (2)	—	—	—	—	—	—	5	12	61	—	—	—
06 Oct (2)	—	—	—	—	—	—	—	—	—	3	15	33
07 Oct (4)	—	—	—	—	—	—	—	—	—	6	38	56
08 Oct (4)	—	—	—	—	—	—	—	—	—	3	24	40
09 Oct (2)	—	—	—	—	—	—	—	—	—	1	8	22
Total	25	161	702	45	249	934	46	222	726	13	85	151

Table 1. Number of GNSS RO profiles assimilated for the Control_noC2 (noC2), Control (CTRL), and ROMEX experiments in the 12° × 12° HAFS-A moving nest for each day that overlaps a TC case cycling period. Dashes indicate that no HAFS DA cycles occurred for that storm on that day, as identified by the column and row respectively. Numbers in parentheses list the number of DA cycles for a given day, with the maximum value of four indicating that observations were assimilated in the 00, 06, 12, and 18 UTC cycles. The comma-delimited values for 23 and 24 September indicate that the HAFS-A nest assimilated profiles for both the Fiona and Ian cases, with the Fiona cycling period ending at 18 UTC 24 Sep and the Ian cycling period beginning at 06 UTC 23 Sep. The sets of profiles assimilated for Fiona and Ian on 23-24 Sep were different and non-overlapping due to differences in the HAFS-A nest position for the two storms.

5. The mechanistic interpretation presented in lines 599–608 is currently somewhat stronger than what is directly supported by the diagnostics shown. Figures 11–13 support the inference that ROMEX analyses are drier in Fiona's inner-core / near-storm environment and that the subsequent forecasts exhibit a weaker inner-core structure and reduced over-intensification bias. By contrast, the proposed links involving enhanced AAM convergence, vertical AAM transport, and vorticity tilting are physically plausible, but

they are not explicitly diagnosed by the present figures. I therefore recommend that the authors either (i) substantially soften this discussion and frame it as a qualitative interpretation consistent with prior theory, or (ii) provide additional momentum or vorticity budget diagnostics to more directly support these claims.

We decided to soften the mechanistic discussion in this passage, given that we have not performed the vorticity and AAM diagnostics needed to support statements linking the commercial RO DA-impacted analysis water vapor fields to ROMEX-versus-Control intensity/structural differences in the HAFS Fiona forecasts. Fiona presents an interesting forecast case study, and we think that it would be best to save the vorticity/AAM diagnostics of our HAFS ROMEX experiments for a future case study that would combine RO forecast impact assessment with deeper scientific analysis. Trying to combine an expanded mechanistic discussion with the statistical DA and forecast impacts presented in Sections 3 and 4 would make this manuscript too long, in our opinion. We revised this passage to emphasize consistency of our results with TC intensification mechanisms rather than make speculative statements that cannot be proven with the results shown here. Starting from the beginning of the paragraph, this section now reads as (lines 670-694), with the revised portions highlighted in bold:

"Figure 13 compares Hurricane Fiona's structure in 12-h Control and ROMEX forecasts initialized at 12 UTC 21 September. While both the Control- (Fig. 13a) and ROMEX-forecast (Fig. 13b) storms' convective patterns share many characteristics, including a prominent rainband located in the northeastern quadrant around 2 degrees' radius from the center, the Control storm has higher reflectivity values within its eyewall. The latter appears here as the annular-shaped region around Fiona's center where 900-hPa horizontal wind speeds exceed 50 m s^{-1}. Compared to ROMEX, the Control forecast's near-surface eyewall winds are about 5-10 m s^{-1} stronger, broadly consistent with these forecasts' P_{MIN} differences (compare the dotted and solid orange lines at 00 UTC 22 September in Fig. 11a) under an empirical TC vortex pressure-wind relationship (Chavas et al. 2025). Control also generates stronger and more organized midlevel updrafts, arcing around the ~ 0.3-degree radius in the northern and southeastern quadrants of its eyewall (cf. 13c, 13d). These differences in reflectivity and mid-level updraft strength between the Control and ROMEX 12-h forecasts reveal that the Control has more vigorous inner-core convection, which aligns with its analysis showing a moister lower-to-middle troposphere within a 100-km radius (Fig. 12). **This is dynamically significant: deep convective updrafts in a TC inner core provide a key mechanistic link between the favorable thermodynamic conditions of high relative humidity (RH) in the lower-to-middle troposphere and the storm's wind field intensification. For example, it is well known that eyewall updrafts facilitate latent heat conversion to kinetic energy (Rotunno and Emanuel 1987) and convergence of higher absolute angular momentum (AAM)-air advected inward from larger radii (Zhang et al. 2001; Montgomery and Smith 2014).** A moister near-storm environment limits dry air entrainment and favors stronger, more organized eyewall updrafts (Cram et al. 2007; Braun et al. 2012). Therefore, it is **plausible** that the drying impact of the assimilated ROMEX observations on Hurricane Fiona's 21 September analyses helped to reduce the storm's over-intensification bias in HAFS-A forecasts initialized from those analyses."

6. Based on Figs. 11–14, the proposed mechanism after line 627 appears physically plausible, but it is not directly demonstrated by the diagnostics shown. In particular, the

inferences in lines 631–637 could be further supported by additional diagnostics, such as vorticity or angular-momentum budget analyses, and warm-core tilt diagnostics, if the authors wish to maintain a mechanistic interpretation at this level of specificity.

We agree that the portion of the submitted manuscript between lines 631 and 637 contains speculative statements about vertical wind shear (VWS) impacts on Fiona that would require additional diagnostic analysis to be supported. Given the complexity of this analysis, we think that it is best saved for another study. Therefore, we have re-written these sentences to emphasize broader consistency with current scientific understanding without delving into any details about how the VWS could have affected Fiona's vortex. We still think that it's worth briefly commenting on how the similarity between the Control and ROMEX intensity forecasts during Fiona's later steady-state and weakening periods are consistent with greater sensitivity to environmental conditions not impacted by RO DA in HAFS, as opposed to the earlier forecast periods when near-storm environmental water vapor likely has a stronger control on the storm's intensity. In our revised manuscript, this paragraph reads as (lines 706-726), with the most heavily revised portion highlighted in bold:

"Revisiting these Control and ROMEX forecasts thirty-six hours later at 12 UTC 23 September, we find more similar intensities, both in terms of their 900-hPa wind speed, which exceeds 60 m s^{-1} in the eastern quadrant (Figs. 14a and 14b), and in their ~ 930 hPa P_{MIN} (orange lines in Fig. 11a). Compared to the 12-h forecast, Fiona's inner-core reflectivity, horizontal wind (Figs. 14a, 14b), and updraft (Figs. 14c, 14d) structures are more asymmetric. Meanwhile, southwesterly 850-200 hPa VWS ahead of an upper-level trough moving off the northeastern US coast has strengthened from 11.0 m s^{-1} to 12.9 m s^{-1} over the past 36 hours, according to the NHC's Statistical Hurricane Intensity Prediction Scheme (SHIPS; DeMaria et al. 2005) analysis. Differential advection of TC eyewall vorticity with height caused by moderate-to-strong VWS creates a more favorable (unfavorable) environment for deep convection on the vortex's downshear (upshear) side (DeMaria 1996; Jones 1995; Chen et al. 2018), which could account for the asymmetries shown in Fig. 14. **Broadly consistent with the Best Track, both the Control and ROMEX forecasts initialized at 12 UTC 21 Sep maintain Fiona's intensity in a quasi-steady state for the next 24 hours, after which the storm weakens (Fig. 11a) while undergoing extratropical transition (Pasch et al. 2023). Although beyond the scope of this study, we speculate that Fiona's weakening resulted from a combination of cooling SSTs and the disruptive impacts of moderate-to-strong VWS on the vortex structure (Simpson and Riehl 1958; Frank and Ritchie 2001). Predominating influences from these larger-scale environmental conditions, which are not affected by DA in the HAFS-A nest (Section 2b), would be consistent with the relatively small differences between the Control and ROMEX intensity forecasts during this later period of Fiona's lifecycle (Fig. 11).**"

7. More broadly, Section 5 contains an interesting and potentially important case study, but the mechanistic interpretation is currently stronger than what the diagnostics directly support. I therefore recommend that the authors either soften the mechanistic discussion throughout this section or provide additional diagnostics to more directly support the proposed dynamical interpretations.

We have softened the mechanistic discussion of Section 5 and re-worded it to emphasize how the RO impacts on Fiona's structure/intensity are broadly consistent with our current scientific understanding of how a TC vortex responds to environmental water vapor, rather than make speculative statements that we cannot support with diagnostics shown in this study. We may consider using our ROMEX and Control experiment outputs for the Fiona case in a future study that digs deeper into studying the physical mechanisms affected by the commercial RO data assimilation.

Technical corrections

1. In the first paragraph, the discussion abruptly shifts from bending-angle retrieval to the roles of water vapor, temperature, and pressure in refractivity. A smoother transition would improve the logical flow.

We moved the clause starting with "since atmospheric refractivity depends on …" from the beginning to the end of the sentence to improve the transition. This passage now reads as (lines 71-78):

"Processing centers, such as the University Corporation for Atmospheric Research (UCAR)'s COSMIC Data and Archive Center (CDAAC), use physics-based inversion techniques (Kuo et al. 2004) and precisely determined transmitter and receiver satellite positions to retrieve vertical profiles of the RO ray bending angle from Doppler-shifted occultation radio signals. RO bending angles can also be simulated from numerical weather prediction (NWP) model output using a forward operator that integrates atmospheric refractivity field gradients along an approximation of the RO ray path (Healy et al. 2007; Cucurull et al. 2013), since atmospheric refractivity depends on water vapor, temperature, and pressure."

2. Line 192: "VMAX stratifies hurricane intensity   according to the Saffir-Simpson Scale". Please remove the extra space in the sentence.

Fixed (line 208).

3. Line 219: Please spell out the full name of "TCVitals" when it is first introduced.

Fixed (lines 237-238).

4. The manuscript uses both "HAFS" and "HAFS-A" in different places. Please make the terminology consistent throughout the manuscript.

We revised the manuscript to use HAFS-A instead of HAFS, and we added additional information when introducing HAFS-A in the Introduction (lines 152-154):

"Together with its "B" configuration, which uses different physics parameterizations and horizontal grid resolution, HAFS-A has provided NOAA with operational regional TC forecasting guidance since 2023."