

Below, reviewers' comments are shown in black, and our responses are provided in blue.

Reviewer #3: This manuscript “ Impact of high-volume GNSS Radio Occultation data on improving the accuracy of Navy's global weather prediction models.” by Christophersen, Ruston and Tyndall assesses the impact of high-volume GNSS Radio Occultation data during ROMEX on improving the accuracy of Navy's global weather prediction models. I recommend to publish this work with some minor revisions, see listed below.

In general, it would be good to see that not only RMSE is used to assess forecast scores but also statistical measures as well, as e.g. std deviation or anomaly correlation.

We appreciate the reviewer's thoughtful comments and suggestions. We agree that it is generally beneficial to consider additional metrics beyond RMSE, such as standard deviation or anomaly correlation, when evaluating forecast performance. In our routine analyses, we do generate these supplementary performance metrics. However, for the purposes of this manuscript, we chose to focus on RMSE statistics, as we believe they effectively convey the key findings regarding forecast accuracy. Thank you again for your valuable feedback.

P3, eq (1); for h lower equals 10km which correction is used at 0km? Does it start with 2% and the linearly decreasing to 0.2% at 10km?

Yes, the threshold value at 0km is 2%, and at 10km is 0.2%, with a linear interpolation for 0-10km.

We added the starting value for $h \leq 10 \text{ km}$ range.

linearly ramped from 0.2% to 2%, $h \leq 10 \text{ km}$

P4, L.110/111: Be clearer what you mean with “normalized by error”, as you normalise background in Fig. 2a and by obs error in Fig. 2b.

We revised the sentence to differentiate the two normalizations for Fig. 2.

P5, Fig.2: Why didn't you plot romex_cntl? It would have been interesting to see how it looks for this experiment. Also colour choice is not colourblind friendly (also applies to Fig.3).

Thanks for the suggestion. We added the control experiment in Fig. 2 and updated all the colors in Fig. 2, Fig. 3, Fig. 6, Fig. 8, and Fig. 10 to be colorblind friendly.

P8, Fig 4. COSMIC-2 satellites naming by E1 to E6 needs to be clearer for the reader. Maybe COSMIC-2 E1....

Added.

P 9, l.212 Could you mention the magnitude of this bias for GP?

The bias for geopotential is about 5 meters. Added in the revision.

P 15, chapter 4. Please mention that also other studies using a more advanced forward operator - e.g. 2D ROPP version – could also be tested if this helps to improve forecast scores, especially in the troposphere.

Added.