

1st review of “Processing Multiple GNSS RO Data Using FSI and ROPP: Results from the ROMEX” by Chen et al.

This manuscript evaluates the performance of the STAR RFSI processing system in deriving atmospheric profiles from level-1b data across multiple GNSS RO missions collected during the ROMEX project. By comparing the results against ROPP (using the CT2 method) and EUMETSAT-processed datasets, the authors further discuss the structural uncertainties inherent in RO retrievals. While the study presents noteworthy results, certain areas remain ambiguous (listed below) and would benefit from further clarification.

Reply:

Thank you for your constructive suggestions, which have helped improve the quality and clarity of this manuscript.

General comments:

1. The FSI method was introduced two decades ago and is recognized as a powerful WO method for resolving fine-scale structures and multipath effects. However, its requirement for specific geometric conditions (e.g. circular orbits) meant that it required significant/complicated correction (as described in Line 363-374) before it could be reliably used in operational RO data processing. In addition, the FSI method is highly sensitive to noise, as acknowledged by the authors. In the lower troposphere, where SNR is often low, FSI performance can degrade significantly. These limitations have likely contributed to the fact that FSI has not been widely adopted in operational RO processing centers. Given these considerations, the manuscript would benefit from a clearer explanation of the scientific or practical motivation for reprocessing data using FSI. Beyond generating an additional version of RO products, what specific advantages does the FSI offer compared to existing operational algorithms?

Reply:

We agree that a clearer motivation is needed. We added a statement to the Introduction emphasizing that “FSI is theoretically well suited for resolving fine-scale atmospheric structures and handling multipath in the lower troposphere (Jensen et al., 2003; Adhikari et al., 2021), a key source of uncertainty for NWP. The use of the STAR RFSI algorithm within ROMEX is therefore to quantify the structural uncertainty associated with this alternative, high-resolution retrieval approach relative to community-standard methods, thereby providing critical insight for optimizing multi-mission data assimilation strategies.”

2. The authors devote a substantial portion of Section 3 to describing the methodology. It would be helpful if they could first clearly summarize the core methodological differences between the approach presented here and that described in Adhikari et al. [2021]. Such clarification would allow readers to better understand the novelty and specific contributions of the current work.

Reply:

The core FSI algorithm remains consistent with that described in Adhikari et al. (2021). The novelty of the present work lies in the development and presentation of the complete STAR RFSI end-to-end processing framework (including data preprocessing, quality control, and statistical optimization) and its systematic, harmonized application to the large and diverse multi-mission ROMEX dataset for comprehensive assessment of structural uncertainties against other

major processing centers (such as EUMETSAT and UCAR). We added these sentences to the text.

3. The title of Section 3, "Full Spectrum Inversion Algorithm", appears somewhat misleading. Based on the content, the section describes the STAR FSI-based processing framework as a whole, including the full processing chain for retrieval Level 2 data products from Level 1b data. This scope extends well beyond the pure FSI algorithm itself, which is fundamentally a WO method for resolving multipath effects in the troposphere. The title could therefore be reconsidered to better reflect the broader processing steps being presented.

Reply:

We agree. We changed the title of Section 3 to "STAR RFSI Algorithm and Processing Chain" to better reflect the broader scope, which covers the full retrieval process from Level 1b to Level 2 data products.

4. The manuscript also introduces several technical terms and abbreviations that are not sufficiently explained or referenced, making it difficult for readers to follow. For example: what're the differences between STAR RFSI and STAR ROPP? Is it correct to simply understand the STAR RFSI uses the FSI algorithm, whereas STAR ROPP applies the CT2 algorithm? What is the core difference between STAR ROPP and the well-known ROM SAF ROPP processing system? Is the STAR ROPP considered the "community standard" ROPP referred to in Lines 19-20 and in the title?

Reply:

STAR RFSI is our algorithm that uses the FSI method to retrieve bending angles. STAR ROPP is the customized ROPP v10.0 system at STAR capable of to running either the CT2 or FSI (RFSI) method. The official "community standard" ROPP is the ROM SAF ROPP, from which our STAR ROPP is distinguished (Lines 77–78). In this study, the community-standard dataset refers specifically to data generated using the STAR ROPP CT2 method. We will explicitly state this in the text.

5. Does EUMETSAT use ROM SAF ROPP operationally to process RO data? If not, what kind of WO method is used in their processing? CT2?

Reply:

For the ROMEX dataset, the EUMETSAT-processed data (except for COSMIC-2) are generated using ROM SAF ROPP as the internal processing system. The COSMIC-2 data in the EUMETSAT dataset were processed by UCAR CDAAC using the Phase-Matching (PM) method (Line 681).

6. Regarding QC (Section 3.4), the authors use ERA5 forecasts as reference for QC checks, and ERA5 reanalysis as a reference for evaluating RO retrievals. Could the authors assess the impact of using ERA5 as a QC reference on the resulting dataset? For example, how sensitive are the retained profiles to the choice of ERA5-based thresholds?

Reply:

The impact of small-scale ERA5 errors on the QC is minimized by employing a very high threshold $7\sigma_{year}$, which is intentionally conservative. This threshold is chosen to filter only the most extreme outliers, ensuring the retained profiles remain largely insensitive to fine-scale variations in the ERA5 forecast field. The resulting QC pass rates are comparable to those from CDAAC-processed data with internal QC procedures. The QC rejection rate for the ROMEX dataset depends the mission. For example, the QC pass rates are 85.7%, 94.8%, and 93.9% for COSMIC-2, Spire, and PlanetiQ, respectively, compared to 87.6%, 95.7%, and 90.6%, for the corresponding CDAAC-processed datasets.

7. In Section 4, global RO retrievals from different ROMEX missions processed by different systems are compared. This raises concerns of the impact of significant sampling differences on the statistics shown in Fig. 5-10. In particular, C2, as the only low-inclination mission, with observations concentrated in mid- or low latitudes, can give significantly different statistical results. In addition, it would be very informative if the authors could include penetration depth statistics for different missions and processing systems in Fig. 5-10. Such information would help readers better interpret differences in lower-tropospheric performance and assess the practical impact of the different algorithms.

Reply:

We agree that sampling differences across missions can influence the statistical comparisons. We have addressed this concern by explicitly discussing the distinct characteristics of COSMIC-2, including its low-inclination orbit and concentration of observations in the tropical/subtropical regions, which contribute to its higher variability and positive bias (Lines 522–527). To further account for sampling differences, we provided additional analysis in Section 5 (Fig. 16), where the comparison is restricted to a common latitude range ($\pm 45^\circ$) consistent with COSMIC-2 coverage. This allows for a more balanced inter-mission comparison.

In response to the reviewer’s suggestion, Figures 5-10 have been revised to adopt a more distinct color scheme, that clearly distinguishes among missions and processing systems. In addition, profile penetration depths below 8 km have been included in Figures 5-7 for different missions and processing systems to aid interpretation of lower-tropospheric performance. As the penetration depth exhibits a similar behavior for refractivity, it is not shown in Figures 8-10.

8: I found the results presented in Fig. 16b particularly interesting, as they show strong agreement between the STAR RFSI-processed C2, Spire and PlanetiQ datasets. This consistency is noteworthy given the findings in Anthes et al. (2025) (<https://doi.org/10.5194/amt-18-6997-2025>), which discuss the bending angle biases associated with the sideways sliding of the tangent point. That research suggests such sliding can introduce a slightly positive BA bias in C2 and typically negative biases in other high-inclination missions. I would appreciate the authors’ interpretation of why the RFSI framework yields such high agreement across different missions.

Reply:

We thank the reviewer for this insightful comment. We agree that the strong agreement among the STAR RFSI-processed COSMIC-2, Spire, and PlanetiQ datasets shown in Fig. 16b is noteworthy, particularly in light of the findings of Anthes et al. (2025), which highlight mission-dependent bending angle differences associated with the sideways sliding of the tangent point.

While the STAR RFSI results exhibit improved cross-mission consistency, the underlying causes of this agreement are still under active investigation (Line 785). One contributing factor may be the treatment of the horizontal sliding of the tangent point within the RFSI framework. The differences in the definition of the occultation point (or georeferencing) can influence the magnitude of the sliding-related correction. As discussed in Anthes et al. (2025), UCAR defines the occultation point based on the location where the L1 excess phase exceeds 500 m, typically in the lower troposphere, whereas ROPP defines it as the location where the straight-line tangent altitude is reached (height of straight line equals 0), typically in the upper troposphere–lower stratosphere (UTLS). This sideways sliding effect can introduce a positive bias in COSMIC-2 bending angles of up to ~0.05%.

We added a brief discussion of these factors in the manuscript to provide context for the observed cross-mission consistency in the STAR RFSI results.

Specific comments:

1. L57-58: Can the authors be more specific about the "particular strength" of FSI, compared to other WO methods in resolving fine-scale structures?

Reply:

FSI relies on explicit signal localization to enhance vertical resolution and mitigate multipath by reducing spectral mixing, whereas CT2 achieves an implicit, physics-based localization in impact parameter space, enabling more robust separation of multipath contributions. As a result, FSI is generally more sensitive to fine-scale atmospheric structures but also more noise-sensitive, while CT2 provides more stable retrievals in strong multipath conditions at the expense of reduced small-scale resolution.

2. L97: “ROEMX” -> “ROMEX”

Reply:

Agreed. Corrected.

3. L113-114: As far as I know, C2, Spire and PlanetiQ are all using OL tracking, instead of the combination of OL and CL tracking. Can the authors recheck the fact and may readdress the statement?

Reply:

We have rephrased the statement to ensure accuracy regarding the latest tracking strategies of these missions. “However, recent GNSS radio occultation missions, including COSMIC-2, Spire, and PlanetiQ, primarily employ OL tracking throughout the occultation in order to maximize tracking robustness and data continuity across all atmospheric layers.” The general statement about the benefit of combining OL and CL tracking will be maintained as a conceptual point.

4. L170: “GNS” -> “GNSS”

Reply:

Agreed. Corrected.

5. L171-172: "STAR RFSI algorithm has been integrated into ROPP version 10.0": Do you mean the STAR RFSI algorithm has been included in the ROMSAF ROPP package? If not, please rephrase the sentence. It's misleading.

Reply:

Agreed. We rephrased to clarify that "the STAR RFSI algorithm has been integrated into the ROPP version 10.0 customized at NOAA STAR."

6. L214-217: Grammatical error. Rephrase it.

Reply:

Agreed. The sentence was rephrased, the grammatical error was corrected.

7. Section 3.2.1: I had thought that the navigation data modulation had been removed in the provided Level1b data, isn't it true?

Reply:

While Level 1b data providers typically remove the navigation bit, it is a best practice to correct for residual navigation bit jumps that may remain in the excess phase data. The model phase serves as a robust and necessary reference for identifying these $\pm\pi$ discontinuities (Lines 292–296).

8. The current presentation of Fig. 4 is somewhat confusing, as different months are represented by colors that appear identical.

Reply:

Agreed. Figure 4 is revised with a more distinct color scheme to clearly differentiate the monthly standard deviations.

9. L465-466: So the RO data may be flagged as "bad" solely because "the model simulation data are unavailable"? This doesn't sound like a good QC criterion.

Reply:

The profile is not being judged as "bad" in the physical sense. The QC relies on the Observed-minus-Background (O-B) comparison. If the Background (B) simulation data is unavailable, the QC check cannot be performed, and the profile is conservatively flagged to ensure that only profiles fully validated against the ERA5 reference are retained. This condition has been removed to avoid potential confusion in the text.

10. L743-745: could the authors elaborate on the statement regarding "degraded L2 signals and reduced GNSS SNR above ~20 km"?

Reply:

We thank the reviewer for pointing this out. There is a typo in the original text: "above" should be clarified to refer specifically to the ~20–25 km altitude range, and we will revise the wording

accordingly. The L2 signal is inherently weaker than L1 and therefore more susceptible to noise. In the 20–25 km region, although the neutral atmospheric signal is already relatively small, the L2 signal quality can degrade due to its lower SNR. This leads to increased uncertainty in the ionospheric correction, which relies on the combined use of L1 and L2 measurements. As the FSI technique is more sensitive to measurement noise, these effects are amplified, resulting in degraded retrieval quality in this altitude range.