Peer Review Report

Manuscript title: Local departures from IRI-2026 and observation-based correction at Hainan Fuke
Authors: Zheng Wang[1,2], Ziyue Wang[1], Jinhui Cai[3,*], Guojun Wang[1,2], Chu Qiu[3], Jiankui Shi[1], Dabin Xue[1], Pengdong Gao[3], and Xiao Wang[1]

Journal: Annales Geophysicae

A. Overall evaluation

The manuscript evaluates IRI-2026 against ionosonde observations at Hainan Fuke, China, during 2014–2020. It examines local model departures in foF2 and hmF2 and compares three correction approaches, including month-hour correction, random forest, and preceding-residual correction. The analysis of correction performance under different spread-F conditions is also interesting.
The topic is relevant to ionospheric modelling, and the use of an independent evaluation period is a strength. The manuscript presents useful observational results and has potential scientific value. However, some methodological issues, particularly data reproducibility, observational uncertainty, spread-F classification, and machine-learning validation, need to be addressed before publication.

Assessment against the journal's criteria

1. Does the paper contain new data or new ideas, or both?
Yes. The study provides locally curated ionosonde observations and evaluates IRI-2026 using different correction methods. The spread-F-dependent analysis adds further value. However, the authors should explain more clearly how this work advances previous local IRI evaluation studies.

2. Are these up to international standards?
Partially. The research approach is relevant, but improvements in data accessibility, quality control, uncertainty assessment, and methodological reproducibility are needed.

3. Is the presentation clear?
Generally, yes. The manuscript is logically organized, although some methodological details, particularly the reconstruction of the quiet-map foF2 baseline and the treatment of storm corrections, require clearer explanations.

4. Does the author reach substantial conclusions?
The conclusions are mostly supported by the results. However, they should be limited to the investigated station and period, since the transferability of the correction methods to other locations has not been demonstrated.

5. Is the length of the paper adequate?
Yes. The length is appropriate for the scope of the study.

6. Is the language fluent and precise?
Generally, yes. Minor language editing would improve clarity and readability.

7. Are the title and abstract pertinent and understandable?

Yes. Both reflect the main objectives and findings. However, the authors should explain the difference between the foF2 MAE values of 1.137 MHz in the abstract and 1.130 MHz in Table 1.

8. Is the size of each figure adequate?
The figures are relevant, but the authors should ensure that labels, legends, and symbols remain readable in the final publication.

9. Does the author give proper credit to related work and clearly indicate their own contribution?
Relevant studies are cited. However, the authors should better distinguish the novelty of this study from previous local IRI evaluations, particularly regarding IRI-2026 and the proposed correction methods.

B. Major scientific and technical comments
1. Data availability and reproducibility
The manually curated ionosonde observations, individual peak records, and spread-F labels are not publicly available. The available code also appears insufficient to reproduce the complete analysis. The authors should provide access to the processed data and relevant analysis code, or explain how independent researchers can obtain them.

2. Uncertainty in ionosonde observations
The study relies on manually reviewed ionograms to obtain foF2 and hmF2. However, the uncertainty associated with manual scaling, particularly during spread-F conditions, is not adequately discussed. The authors should explain the quality-control procedure and provide an estimate of measurement uncertainty.

3. Spread-F classification

The reliability of the five spread-F classes needs further clarification. The authors should provide information about the validation and accuracy of the classification method and discuss the possible influence of classification errors, especially in the smaller strong-range-spread group.

4. Random forest validation
The random forest model uses 2015–2018 for training, 2019 for validation, and 2020 for evaluation. Since year is included as a predictor, the authors should explain how the model handles prediction beyond the training period. They should also clarify the purpose of the 2019 validation dataset.

5. Comparison with a simple persistence method
The preceding-residual correction should be compared with a simple persistence approach:

$$\hat{y}(t) = y(t - \Delta t)$$

This would help establish whether the proposed method provides an improvement beyond simply using the previous observation. The authors should also clarify whether the preceding observation is available in real time.

6. IRI-2026 implementation and storm correction
The reconstruction of the quiet-map foF2 baseline from storm-adjusted NmF2 requires a clearer mathematical

explanation. The authors should provide the exact equation, define the storm multiplier, and document the relevant IRI-2026 settings to improve reproducibility.

7. Data quality and statistical reliability

The authors should justify retaining physically unusual observations and provide a sensitivity analysis showing whether these values affect the main results. In addition, the relatively small sample sizes in some spread-F groups should be considered when interpreting the statistical improvements.

8. VTEC comparison and study limitations

The authors should discuss the differences between GNSS-derived VTEC and IRI-integrated VTEC, particularly their integration limits and measurement uncertainties. Furthermore, because the study uses only one station, the proposed corrections should be described as local approaches rather than generally applicable improvements to IRI-2026.

C. Minor comments

i. Clarify the difference between the evaluation samples in Tables 1 and 2.
ii. Define all variables clearly and check the formatting of equations.
iii. Report the number of observations retained after quality control.
iv. Explain the treatment of missing predictor values
v. Discuss possible correlations among the random forest input variables.
vi. Ensure that figure captions and spread-F abbreviations are consistent.
vii. Check the reference list for formatting and DOI consistency.

D. Overall recommendation

Major revision

The manuscript addresses an important topic and presents useful observational results on the local performance of IRI-2026. The comparison of different correction methods and the analysis of spread-F conditions are valuable aspects of the work. However, the issues related to data availability, observational uncertainty, classification reliability, and machine-learning validation need to be addressed to strengthen the scientific reliability and reproducibility of the study. I recommend major revision before the manuscript can be considered for publication.

E. How would you rate the scientific contribution of the manuscript?

May have potential after additional work and resubmission.

The study has scientific potential, but further methodological clarification and validation are needed to establish the robustness and broader significance of its findings.