General comments

The manuscript presents an evaluation of the WRF-Chem model in simulating chemical species over Réunion Island, a remote Tropical island. The authors explore various model refinements aimed at improving the performance of WRF-Chem in predicting key species such as formaldehyde, methanol, isoprene, Iox (isoprene oxidation products), monoterpenes, acetone, NO2, NOx, OH, and O3. Overall, this manuscript presents valuable work on the evaluation of the WRF-Chem model in simulating VOC and NOx species in a tropical environment. However, it needs significant improvement in terms of clearly defining scientific questions, providing statistical evaluation of the model performance on both meteorological and chemical species simulations. I recommend resubmission after these revisions.

Major Comments:

1. The primary weakness of the manuscript is the absence of clearly defined scientific questions. While the study provides an extensive model evaluation and sensitivity analysis, it is not entirely clear what the specific scientific questions are that the authors aim to address. The manuscript seems more focused on optimizing model performance and testing various sensitivity runs rather than contributing new scientific insights. I suggest that the authors clearly articulate the overarching scientific objectives of their study, and emphasize how the model evaluation contributes to advancing our understanding of VOC and NOx budgets in remote tropical regions.

2. The manuscript provides a significant amount of detail on the instruments and data collection methods, including extensive descriptions of the Proton Transfer Reaction Mass Spectrometry (PTR-MS) measurements and other observational data sources. However, the focus of the paper is on evaluating the WRF-Chem model's performance, not on collecting these VOC data. The detailed description of the instruments and data collection could be condensed to make the manuscript more focused. The authors should focus on the model evaluation and sensitivity tests, while providing only essential information on the data sources used in the model evaluation.

3. While these adjustments are important for improving model performance, the manuscript does not clearly highlight the scientific significance of these changes. For example, the adjustment of power plant NOx emissions by a factor of 5 is a major model refinement, yet the scientific reasoning behind this adjustment and its broader implications for understanding VOC and NOx budgets in tropical regions should be more clearly articulated. The manuscript could benefit from a more focused discussion on the implications of these model refinements for improving our understanding of atmospheric chemistry in tropical regions, particularly in terms of source attribution and budget constraints.

4. The manuscript provides a thorough comparison of model results with observational data, including comparisons with PTR-MS measurements and satellite data (e.g., TROPOMI). While the model generally reproduces the observed VOC and NOx concentrations reasonably well, there are significant discrepancies, such as overestimations of surface ozone and underestimations of certain VOC species (e.g., formaldehyde). These discrepancies should be discussed more comprehensively, with a focus on their potential causes and implications for the model’s representation of tropical atmospheric chemistry. Furthermore, the authors should consider including a more detailed discussion on how these model performance issues could be addressed in future work, especially with respect to the neglected halogen chemistry and potential missing sources of key compounds.

Specific comments:

1. As indicated in Lines 158–160, the simulations were conducted sequentially in 2-day intervals. However, this approach is not appropriate for completing one-month-long simulations. It is recommended to perform 48-hour (2-day) simulations daily, starting at the same time (e.g., 00Z or 12Z UTC). Evaluations should focus on the results from either the first 24 hours (Day 1) or the second 24 hours (Day 2) of each simulation, rather than combining sequential 2-day simulations into one-month simulations. This distinction is crucial because numerical models perform differently at various forecast hours. The authors should exercise caution to ensure the model performance is evaluated realistically.
2. Please simplify the description of biogenic emissions in Section 2.3.2, unless any updates or changes have been made to MAGAN 2.0.4. For example, is it necessary to retain the content from Lines 302 to 324?
3. It is recommended to simply the description of observational data sine you are the data users rather than the data collectors.
4. I understand, this is what you have for model evaluation but one-site evaluation is not suffice to assess the model performance. Meanwhile, it would be better to provide calculation of statistical parameters (e.g., correlation coefficient, root mean square error, mean bias etc.) for model evaluations.
5. Over a large portion of the island (southwestern part as well as a small portion of the northern coast) but also over a large oceanic region surrounding almost the entire island (especially to the northwest due to the influence of Le Port emissions), OH levels are enhanced, primarily because NOx promotes the conversion of HO2 to OH (Spivakovsky et al., 2000) through the reaction r5. If this is true, should we see enhanced NO2 but we do not see that in Figure 9 a and e panel.
6. Over large portions of the island (southwestern areas and parts of the northern coast) and surrounding oceanic regions, OH levels are enhanced due to NOx-promoted conversion of HO2 to OH (Spivakovsky et al., 2000, reaction r5). If this explanation is valid, enhanced NO2 levels should also be observed, but this is not evident in Figure 9 (panels a and e). Please clarify.
7. Figure 10: Please include a comparison of simulated and observed O3 at the Le Port station to assess the impact of NOx and VOC predictions on surface O3 levels.
8. Lines 368–369 indicate that there are 18 air quality monitoring stations. It would be more appropriate to include all 18 sites in the evaluation by calculating statistical parameters such as the correlation coefficient, RMSE, IOA, and others.

Minor comments

1. Line 23: Define “a.s.l.”
2. Line 12: Add the location of Piton de la Fournaise to Figure 3.
3. Line 176: Define MEGAN upon first use and include references.
4. Line 190: Spell out full names of MVK and MACR.
5. Line 230: Specify what is meant by “these higher resolution data (0.75° x 0.75°).”
6. Line 233: Spell out the full name of “BIGALK.”
7. Line 241: Spell out "EDGAR" and ensure all abbreviations are defined upon first use.
8. Line 242: Spell out "HTAP" in full.
9. Line 354: Avoid repeating the definition of ISOPOOH (defined on Line 196). Review the document for similar duplicate definitions.
10. Line 560: Define NOx (=NO+NO2) upon its first appearance.
11. Lines 578-579: please provide a list of stations that are included in other stations.
12. Line 582: Correct "cyan" to "blue" and check for similar errors (e.g., Line 588).
13. Line 583: Define the stations referred to in the study.
14. Line 591: No green dots are visible in Figure 2; clarify if they are the red dots.
15. Line 632: Specify the elevation of the Maido station.