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 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.
2. 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.
3. 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?
4. Regarding the sensitivity run S1: As noted by the authors in Lines 598–600, the significant over-prediction of NO2 and NOx at the LP station is primarily due to all emissions being injected at the surface level in the model. A more realistic approach for this run would be to inject most emissions at the plume rise height, rather than reducing NOx emissions from power plants by a factor of 5.
5. Figure 8: Assessing model performance on meteorological simulations based on a single-site comparison is insufficient. Are there additional meteorological observation sites on the island? Additionally, it would be beneficial to include statistical parameters (e.g., correlation coefficient, root mean square error, mean bias, etc.) to enhance the model evaluation.
6. 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.
7. 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.
8. Lines 33-34: It is not appropriate to include the statement “Surface ozone concentrations are overestimated by ~6 ppbv on average, likely due to the neglect of halogen” in the abstract. Although halogen chemistry is not included in the WRF-Chem version used in this study, the overprediction of ozone could result from various factors. These may include uncertainties in anthropogenic emissions, physical processes such as vertical mixing, model resolution (2.5 km is insufficient to resolve detailed wind structures), and atmospheric chemistry. It is therefore inappropriate to attribute the overprediction solely to the neglect of halogen chemistry.
9. While this study made significant efforts to improve WRF-Chem predictions of NOx and VOCs, compare them with multi-platform observations, and assess the impact of various factors through sensitivity studies, it is particularly noteworthy that the authors dedicated substantial effort to updating the chemical mechanisms. These updates are detailed in Sections 2.2.3.1 and 2.2.3.2 (Table 2) and examined in sensitivity run S4. If these updates represent a novel development introduced by this study, they should be explicitly highlighted in the abstract, as they constitute a significant contribution to model development. However, if these updates were derived from previous studies, the authors should clearly specify their sources. Based on my understanding, this study is the first to implement these updates to the chemical mechanisms. If that is the case, the manuscript should be considered as a model development study rather than a model evaluation.
10. It is well established that VOCs and NOx are key precursors in O₃ formation, a point further emphasized by the authors in the abstract (see Line 55). However, as shown in Fig. 18, there are significant differences in the simulated concentrations of key precursors such as HCHO (Fig. 18a, g), CH₃OH (Fig. 18b, h), C₂H₆ (Fig. 18c, i), and PAN (Fig. 18e, k) across different scenarios. Despite these variations, O₃ predictions remain largely consistent across scenarios (Fig. 18f, j). Could the authors explain this discrepancy based on their study or analyses?

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 the stations that are included in the category of “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 given in Figure 2. Do you mean the red dots?
15. Lines 613–614: The authors stated, “This discrepancy between the model agreement for NOx and for NO2 is likely due to the model overestimation of surface O3”. However, NOx is one of precursors to O3 formation, rather than O3 being a precursor or driving factor for NOx predictions. This explanation is not appropriate.
16. Line 629: It should be “in Fig. 14”?