This manuscript outlines attempts to improve the modelling of O_3 concentrations, and the production of NOx-VOC isopleths, by box modelling simulations. The paper has two halves. The first half describes a series of chamber experiments in which a chamber-specific mechanism is produced to improve O_3 concentration predictions. The second half attempts to use the developed chamber-specific mechanism to account for model biases in a box model of the ambient environment, suggesting that heterogeneous reactions could account for over-predictions in ozone concentrations from the base simulation.

The experimental procedures for the chamber experiments seem sound, and while there is some room for improvement in the explanation of the method by which the chamber-specific mechanism rates are selected (as explained later), the produced mechanism appears to be in line with existing literature and reproduces the observations well.

However, I believe that there are numerous issues with the implementation of the ambient box models, including the author's focus on NO_2 uptake as the sole explanation for model discrepancies. Some of these issues may be clarified with additional context in the explanation, but others may require deeper consideration. I have organised my comments into 4 sections: NO_2 Sink, General Comments, Specific Comments, and Minor Comments. I do believe there is value in efforts to improve O_3 predictions and that the potential role of NO_2 uptake is an important consideration, and I hope that addressing these comments can provide the improvements in the manuscript that I believe are necessary for publication.

NO_2 Sink

Generally, I feel that the entire manuscript fails to consider causes for the O_3 measurement discrepancy other than NO_2 uptake. The recommended workflow in Figure 5 completely omits the potential for processes (either chemical or physical) other than NO_2 uptake to influence O_3 concentrations. Given an initial over-prediction in NO_2, it would be possible to tune the NO_2 uptake rate to obtain 'accurate predictions' of any O_3 concentrations. However, this does not mean that the model is accurately representing the chemistry occurring, and the isopleth produced by such a model will not be accurate.

- At Line 300, the authors note that use of their NO_2sink model gives better O_3 predictions. This is only the case because the NO_2 sink has been tuned to give good O_3 predictions. Line 298 states that use of the Basic model distorts the produced isopleth, which is certainly true, but if the NO_2 sink model is "getting the right answer for the wrong reasons", then this isopleth would also be distorted in different ways.

- The authors do not provide any information on physical loss processes represented in the ambient box models. It is typical to include a first-order loss term for all species in the model (dilution term) to prevent the build-up of long-lived species. This term approximates physical removal of species in the real-world due to deposition and transport. Often a simple 24-hour loss lifetime is applied, though ideally a more detailed representation will be used that accounts for changing ventilation and deposition over time (e.g. Whalley et al. 2021 https://doi.org/10.5194/acp-21-2125-2021, Mayhew et al. 2022 https://doi.org/10.5194/acp-22-14783-2022).
  If the authors have not included such physical loss processes, then this is a major oversight that has the potential to explain the observed over-prediction.
  If the authors have included some representation of physical losses, then this needs to be explained in the manuscript. Additionally, the sensitivity of the O_3 concentration predictions to changes in physical losses should be assessed. This would provide evidence that the O_3 over-prediction needs to be explained chemically rather than simply being the result of the representation of physical processes. How does the O_3 concentration change if the physical loss rate is varied within the range of realistic values?
  In addition, if possible, measurements of VOC oxidation products could be compared to modelled concentrations to demonstrate that the model predicts those species well, especially once the large NO_2 sink is added. If the SVR NO_2 sink model predicts both VOC oxidation products and O_3 well, then this would lend further support to the presence of a large NO_2 sink.

- Related to the previous point, I believe that the authors' conclusions regarding the additional NO_2 loss term are far too strong, especially in light of the lack of investigation into the effect of physical processes on the conclusions. For example, Line 279 states that the model results "confirm the existence of unidentified NO_2 sinks". This is inappropriate given the large magnitude of the sink and the lack of physical mechanism offered by the authors. Similarly, Line 319 states that their model "revealed the critical influence of the unidentified NO2 sink".

  - Line 285: The authors state that heterogeneous NO_2 reactions are "likely the major reason" for over-predictions in NO_2. This statement requires justification, especially since the authors' own models demonstrate that adding heterogeneous NO_2 reactions does not correct the bias.

General Comments

- One of the central premises of the paper is to investigate the potential for non-HONO mediated ground reactions to impact atmospheric chemistry (Line 45 and 61). Figure 3 demonstrates that these reactions make an imperceptible difference to the model predictions, which the authors suggest is due to the small surface area-to-volume provided by the ground surface. This seems to be a good conclusion. Given that the introduction section states that the ground-mediated reactions are "systematically unassessed", the authors could highlight this conclusion more (e.g. in the conclusions section). This work seems to suggest that we have been right to ignore these surface processes in previous ambient box modelling studies (except maybe for the case of NO_2/HONO processes).

- Related to the previous point, I believe that it would be useful for the authors to illustrate how much of the additional NO_2 loss in the SVR NO_2Het model comes from the ground reaction vs. aerosol surface reactions.

- As with the representation of physical processes, the authors do not state whether a spin-up period was used for the simulation. The authors should note the length of the spin-up period used in the models, as is common practice.

- Is it true that all of the reactions from Table 2 are implemented in the ambient models (SVR model), with the rates scaled based on the ground surface area? If so, this assumes that the ground surface behaves chemically like the Teflon walls of the reaction chamber, which is a large assumption. The authors should highlight this assumption in the manuscript and discuss the ways that this might not be accurate. For example, Table 2 includes a reaction converting OH to HO_2 which is justified by the off-gassing of organics from the chamber walls. This off-gassing is unlikely to occur in the same way from non-plastic ground surfaces, and any VOCs produced in this manner would be measured during the campaign if they were of a high enough concentration to influence the chemistry.

  Since Figure 3 demonstrates that the inclusion of these ground surface processes makes no difference to the predicted O_3 concentrations, it is not of great importance for the authors to try to refine the wall reactions further, but the assumption of similar chemical behaviour between the chamber walls and the ground must be approached with an appropriate level of caution.

Specific Comments

- Line 87: The authors should explain why the fan was turned off during the experiments. It is my understanding that it is standard practice to leave fans running during experiments to ensure homogeneous mixing of the chamber air. For example, the authors acknowledge that wall processes influence the chemistry in the chamber. If the reaction mixture is not continuously mixed, could there not be potential for concentrations gradients to develop in the chamber air due to these wall processes? Alternatively, if poor mixing could occur as a result of air being removed/added at specific points in the chamber.

- Line 180: The authors note that they calculated the O_3 wall loss rate constant from experiments. They should explain in more detail how this was done. Did they perform experiments in the absence of VOCs and measure the decay rate of O_3? If so, it would be good to see the data from these experiments as supplementary information.

- Line 218: The authors should explain in greater detail why they are using Equation 3 to calculate the surface area to volume ratio, including units for the value of 1.7, explaining what this value denotes (increased surface area from surface roughness etc.), and why they believe it is appropriate to use the Li et al. 2010 value of 1.7 in their simulation (was it a similar ground-type?).

- Line 233: The authors note that they constrained the photolysis rate of NO_2 to measured values but do not state how (or if) they constrained photolysis rates of other species. AtChem2 has a JFAC feature that scales all photolysis rates based on a given measured photolysis rate, was this used here?

- Table 3: Is it the case that the final reaction in this table is the NO2 sink added only in the SVR NO_2sink model? If so, this is not clear as Line 259 mentions Table 3 before discussing the additional NO_2 uptake term. I would suggest either removing this entry from the table and stating the uptake rate constant in the main text (e.g. at line 270) or marking the NO_2 -> product reaction with an asterisk noting that it is only included in the SVR NO_2sink model. Either way, it should be made clear that this reaction is not in the SVR NO_2Het model.

- Line 278: The authors state that the over-predictions observed between the 16th and 17th of August are the result of a decrease in NO_2 during this period. Why is this not captured by the model considering NO is constrained?

- Line 301: The authors' discussion of how NO and NO_2 concentrations change when changing NO_x seems plausible but given the availability of the data, the discussion should be accompanied with evidence from the model. For example, plots of NO to NO_2 ratio or HO_2 and RO_2 concentrations could be constructed.

Minor Comments

- Line 30: The authors should reference the WHO report that states the O_3 exposure limits, rather than linking to a web-page.
- Line 37: The authors state that O_3 formation occurs under "three distinct regimes". The use of the word distinct here is inaccurate. The presence of transitional regimes highlights that the boundary between NO_x-limited and VOC-limited regimes is not clear-cut (i.e. the regimes are not distinct).
- Line 39: The authors reference "the complexities of atmospheric conditions" that make modelling O_3 difficult. This is a vague statement, and the authors should clarify what they mean as there is difficulty in modelling many aspects of O_3 concentrations (emissions, meteorology, and atmospheric chemistry), while the focus of this work is only on chemistry.
- Line 96: The MCM is described as an "explicit mechanism", whereas it should be described as "semi-explicit" since it does perform lumping of some species, most notably in its treatment of RO_2+RO_2 reactions.
- Line 103: The correct reference for the AtChem2 description paper is Sommariva et al. 2020
- Line 224: It would be useful to highlight that while H denotes the boundary layer height, it is being used to calculate the volume of air present above 1 m^2 of surface. I.e. the units of H in Equation 3 are actually m^3, but the volume is equal to the height since the calculation is being performed for 1 m^2.
- Line 227: The values of 2.9E-3 and 5.8E-4 do not line up with the values obtained from Equation 3. I believe that this is just a typo since the correct value of 5.7E-3 is used further down at Line 247.
- Line 227: At the end of this paragraph, the authors mention a "further revised model", which I believe is the model they have been already discussing for the whole paragraph. I think it would be clearer if these final two sentences were moved to the start of the paragraph as an introduction to the section.
- Lines 231-239: Most of this paragraph is describing settings for the model, so should probably be included in the "Experimental Methods" section, as opposed to "Results and Discussion".
- Line 234: Again, I believe that the discussion of why NO_2 and HONO were not constrained would fit better in the "Experimental Methods" section.