

O₃-precursor relationship over multiple patterns of time scale: A case study in Zibo, Shandong Province, China

Zhensen Zheng et al.

This manuscript describes a recent 5-month field campaign to better understand the NO_x-VOC sensitivity of ozone during the summer months in Zibo, China. The authors conducted 0-D box modeling with the MCM v3.3.1 near-explicit mechanism, and determined that the selected time scale for this modeling (i.e. daily, weekly, monthly, or campaign-wide) can affect the magnitude of the dependence of ozone formation on its precursors, with shorter time scales (i.e. daily) leading to a wider range of relative incremental reactivities (RIR). RIRs determine the expected effect of reducing certain emissions on the production of ozone, so are a useful tool in mitigating ozone pollution. The authors determine that the RIRs can vary within a city, due to local emissions, and can also vary within a season. This indicates that care should be taken when ascribing meaning to RIRs, as they can be dependent on the modeling parameters selected.

In general, I think this paper is worth publication in ACP. The main result, that averaging over several months blurs out some relevant chemical complexity, isn't that surprising, but the paper is well written, and it is good to see these results analyzed so completely and clearly. I think it will help the community choose time scales wisely.

I have a few minor suggestions/comments that should be addressed prior to publication:

General comments:

- 1) What is the authors recommendation for time scale? On line 35, they state that “integrating multiple patterns of time scale is useful to derive reliable and robust O₃-precursor relationships”. Do the authors not think that the daily time scale is the most accurate because of its detail? What would be the benefit of doing a weekly or monthly average instead?
- 2) A dilution parameter of 3/86400 s⁻¹ was chosen based on its best fit to the data. Is this the only parameter that was tuned to fit the model to the data? Is it possible that the trends determined here could be affected by the selection of that dilution parameter?
- 3) In general, some of the Supplemental Figures and Tables are not presented in the order in which they appear in the manuscript.
- 4) On line 392, the authors state that the RIR_AVOC (and others) are increasing as the time scale gets narrower. But looking at Figure 6, it seems that only the mean is increasing. The median stays the same, implying that there are some more extreme values of RIR_AVOC in the daily model, that are getting averaged out as the time scale broadens. I think this is a different statement than saying that the RIR_AVOC value changes, and the authors should be more careful about making that distinction.
- 5) Section 3.7 describes the uncertainty analysis, but I believe the authors are conflating the mathematical terms for uncertainty (as analyzed by standard deviation of the averages), and the broader qualitative term for uncertainty (that some chemical species weren't included in the model). Figure 10 does not seem to be informative, it simply demonstrates that more data reduces the uncertainty, but that's well known. I would recommend removing that figure altogether, and focusing section 3.7 on what is missing in the model.
- 6) The RIR_CO is presented in many figures but never discussed. Does CO play a relevant role in the O₃ formation here? If not, why is it presented in these figures?

Specific comments:

Line 19 and 91: “integrating” → “integrates”

Line 21: “multiple-site” → “multiple sites”

Line 26: The authors state that the RIRs are “consistent with time scale”, but the manuscript demonstrates that they have different magnitudes. Do they mean consistent with regard to sign (i.e. all positive or all negative)? This should be stated more clearly. Same comment for line 33, which describes consistency in the photochemical regimes, when I think the authors mean consistency in the sign, but not the magnitude.

Line 58: Describing the 0-D model as “advanced” makes it sound like it is more complex than the regional scale air quality models on line 55. But I would say that the benefit of the 0-D model is its relative simplicity (despite the larger MCM mechanism), which allows the kind of in-depth chemical analysis that the authors do here.

Line 112: The Thermo Scientific 42i measures total NO_x or NO. Can you describe more fully how the measured NO_x was separated into NO and NO₂? Do you expect any uncertainty there to make a difference to your modeling results?

Line 116: More information about the Zibo Eco-Environmental Monitoring Center is needed here. Is this data publicly available? What instrumentation was used? Or is there a reference to this site?

Line 125: “Tenax GR” → “Tenax GR cartridges”

Lines 128 – 137: It would be helpful to describe what differences might be expected between the FID and FID/PID instruments? Have they been cross-checked and calibrated?

Line 139: Were the 5 point calibration standards from 5 separate standard cylinders? Or dilution one cylinder. If the second option, does that affect the accuracy of the calibration at all?

Line 155: Define F0AM

Line 174: “due to significant miss” → “due to significant missing data”. Additionally, which data was missing? It is hard to see in Figure 2 what is missing. Is the cause of the missing data related to the fact that these time periods seem to overlap to unusual (i.e. not very diurnal) patterns in the O₃, such as Jun 3 – 6, Jul 5 – 8, and Aug 9 – 15 in figures S6 - 8?

Line 214: In this analysis, was the 10% change in X a 10% increase or a 10% decrease? Does it matter which is selected, since O₃ has a non-linear response to NO_x in particular?

Line 228: Do the authors have a hypothesis for why the wind speeds were so different at the different sites? Were they at equal altitudes above the ground? Could it have been an instrument issue?

Line 234: How was the scaling done?

Line 243: “quantity” → “quantify”

Line 253: See my general comment #2. Isn't model performance also due to the selection of the dilution parameter that maximizes model performance?

Line 287: Why is there no section dedicated to discussing the full campaign time scale? I only see three sections (monthly, weekly, and daily), but later on the authors discuss the four different time scale analyses.

Line 289 – 295: This is helpful information for understanding how the model was run, and should be moved to the methods section. This is also true for lines 314 – 319.

Line 396: How are these underestimation percentages being calculated?

Line 530: While it is true in this case that all time scales yielded the same information for ranking the top-10 VOC contributors, do the authors expect it would be true in all cases? Unless they have done that analysis, I would recommend changing the language here to just describe these results, and not try to make this broad statement about all models.

Line 546: “difference” → “differences”

Table 1: This is a summary of the most relevant 0-D box models, correct? In that case it should be stated as “Summary of relevant published 0-D box model studies”, so it doesn’t imply this is every box model ever published.

Text S1: This isn’t ever reference in the main text, and doesn’t add any new information. I would delete it.

Line S88: This URL should be a cited reference instead.

Table S2: State in the caption what the asterisk next to “alkenes” refers to. Additionally, does “non-listed in box model” mean that they were measured but not modeled? Or modeled but not measured?