

Atmos. Meas. Tech. Discuss., referee comment RC2
<https://doi.org/10.5194/amt-2021-367-RC2>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on amt-2021-367

Anonymous Referee #2

Referee comment on "Ozone formation sensitivity study using machine learning coupled with the reactivity of volatile organic compound species" by Junlei Zhan et al., Atmos. Meas. Tech. Discuss., <https://doi.org/10.5194/amt-2021-367-RC2>, 2021

This article estimates the sensitivity of ozone formation using a random forest model with not only total VOCs concentrations but also observed concentrations and initial concentrations of VOC species. The result showed that the ozone prediction performance using initial concentrations of VOC species was better than that using total VOCs concentrations.

Analytical reports using machine learning with total VOCs concentrations have been published recently. This article simply indicates the superiority of using overserved or initial concentrations of VOC species. From this aspect, the significance of this study is evident. The reviewer would recommend it for publication.

However, the reliability of analytical data and explanation of the initial concentration of VOC species are not sufficiently indicated in this article. And some expressions seem to be somewhat inadequate for well understanding. Slight revisions are required.

Line 50-75: It is easier to read if you divide the paragraph appropriately, e.g., Line50, Line53, Line59, Line64, Line75.

Line 118-129 and S1: As the composition of VOC species varies greatly from year to year as shown in Fig. 1(F), the analytical reliability is important. Information about VOC measurement in this article is insufficient to understand the reliability. Further information such as observation period and the reason, sampling time or cycle, calibration using standard gas, and method of quality control should be described.

Line 129 and S2: Explanation about PIC is insufficient in this article although PIC is important for the results. To calculate PIC, the initial ratio of Ethylbenzene and xylene must be constant. However, they may be emitted from several sources, e.g. painting, mobile exhaust, etc. Please explain why you can use these compounds in this study. And please describe the VOC sampling time which is also important to calculate PIC, because chemical reactions in the air are different in daytime and nighttime.

Figure 1: What are the red lines in (A)?

Figure 2: It is difficult to find the difference in (A), (B), and (C). Something will be needed to make it clear.

Line 203-205: It is unclear that this is what part about Figure 3A.

Figure 3A: Please explain why RI does not change so much even though the composition of VOCs differs greatly between 2015 and 2016.

Line 287-289: It is unclear why you can describe that the RF model is better than the box model from Figure S5.