Response to Reviewer 2

We would like to thank the reviewer for their detailed edits and comments that have improved this manuscript. Below you will find our response to each comment/question. The reviewer's comments/questions are in bold italic font.

General Comments The manuscript lacks a clear explanation regarding the impacts of its results and how it advances knowledge. In the abstract and introduction, there is discussion of the uncertainties associated with organic acids including how organic acids are commonly not included in detail in models and that few studies on organic acids in the Northeast U.S. exist. The manuscript then states "To address these shortcomings, this study investigates…". How are these shortcomings addressed with the results of this study? This is not clearly explained or highlighted since the conclusions seem to just restate the motivation of the study – there are uncertainties in organic acid production and further modeling and observational studies are needed. While future work can still be needed, it is important to address the significance of the results of this study that was performed and in what particular aspects it furthers current knowledge.

We appreciate the comment, which prompted us to revise the abstract, introduction and conclusions to clarify significance of our study. We hope that we have made it clear that our study highlights the key processes affecting organic acid formation.

The changes are the following:

The abstract was updated to read:

Organic acids represent an important class of compounds in the atmosphere but there is limited research investigating their chemical production, particularly in the Northeast U.S. To improve our understanding of organic acid sources, a modeling analysis was performed for air masses reaching the summit of Whiteface Mountain (WFM), New York where measurements of organic acids in cloud water have been collected. The analysis focuses on a pollution event associated with a heat wave that occurred on 1-2 July, 2018 that exhibited unusually high concentrations of formic (HCOOH), acetic (CH_3COOH), and oxalic (OxAc) acid in cloud water. Gas phase production of organic acids for this pollution event was modeled using a combination of the regional transport model WRF-Chem, which gives information on transport and environmental factors affecting air parcels reaching WFM, and the Lagrangian chemical box model BOXMOX, which allows analysis of chemistry with different chemical mechanisms. Two chemical mechanisms are used in BOXMOX: 1) the Model for Ozone and Related chemical Tracers (MOZART T1), and 2) the Master Chemical Mechanism version 3.3.1 (MCM). The WRF-Chem results show that air parcels sampled during the pollution event at WFM originated in central Missouri, which has strong biogenic emissions of isoprene. Many air parcels were influenced by emissions of nitrogen oxides (NO_x) from the Chicago Metropolitan Area. The gas-phase oxidation of isoprene and its related oxidation products was the major source of HCOOH

and CH_3COOH but both mechanisms substantially underproduced both acids compared to observations. A simple gas+aqueous mechanism was included to investigate the role of aqueous chemistry on organic acid production. Aqueous chemistry did not produce more HCOOH or CH_3COOH, suggesting missing chemical sources of both acids. However this aqueous chemistry was able to explain the elevated concentrations of OxAc. Anthropogenic NO_x emissions from Chicago had little overall impact on the production of all 3 organic acids. Further studies are required to better constrain gas and aqueous production of low molecular weight organic acids.

Lines 34-37 were changed to: "Despite their ubiquity and their growing chemical importance in many regions around the world, organic acids are often not routinely included in studies monitoring the chemical composition of cloud and rain water and are rarely investigated in detail within modeling studies in either the gas or aqueous phase. To contribute to this limited body of work, this study investigates the key processes in both the gas and aqueous phases that led to unusually high concentrations of organic acids measured in Whiteface Mountain (WFM) cloud water on July 1st, 2018.".

Lines 485 – 492 were updated to read: "A large contributing factor to uncertainties in organic acid production is the lack of observational data, particularly organic acids in both the gas and aqueous phases. Regular observational studies over a broader range of geographical and temporal scales are required to better constrain organic acid concentrations. VOC measurements of key organic acid precursors like isoprene, methacrolein, methyl vinyl ketone, and glyoxal, especially in regions of high BVOC emissions, are needed to better constrain organic acid production. Cloud water chemistry measurements must be expanded beyond organic acids to include key aqueous precursor gases such as glyoxal and methylglyoxal. Simultaneous gas and aqueous phase field measurements are also necessary, as cloud water measurements alone are not sufficient to properly investigate cloud water processing of organic carbon. Finally modeling work at different temporal and geographic scales coupled with field observations is necessary for improved representation of organic acids so that the processes governing atmospheric chemistry as a whole. The procedure of back-trajectory analysis that then initialize forward trajectory runs within WRF-Chem (or another chemical transport model) could be automated to provide insight to researchers during field campaigns and guide laboratory analysis of collected samples to target specific chemical species or processes."

The manuscript tends to use qualitative descriptive language and would benefit from more quantitative evaluations. For example, throughout the manuscript there are statements of "good model agreement", "significant reductions", "stronger production", "very little change", "substantially underestimated", "little correlation". How are these terms defined? What quantitative values do they represent? A statement such as "reduced by 80%" or "reduced with statistical significance by t-test" is more informative and stronger than "significant reductions". More quantitative results may also help in addressing the above concern.

The authors agree that we were not quantitative enough while reporting results within the manuscript, particularly within the WRF-Chem evaluations section (Section 3). To address this

comment, we updated the WRF-Chem evaluation section to include new supplemental figures that show linear correlation coefficients and mean bias error (MBE) statistics for the WRF-Chem maps and these same statistics were applied to the WRF-Chem time series plots (now Figures S9) in the supplemental material. Section 3 was updated to read:

"Modeled O_3 exhibited a strong positive linear correlation ($r > 0.8$) with observations across the model domain, but consistently exhibited a mean bias error (MBE) of 10+ ppbv on June 29th and July 1st (Figures S7 and S8). This high bias in O_3 production has been reported in other recent works (Travis et al., 2016; Schwantes et al., 2020; Place et al., 2023) which may be due to overestimated NO_x emissions and/or improper representation of gas-phase organic chemistry. Note that the 2017 EPA NEI used in this study is appropriate for a typical summer day and will likely not represent the actual emissions of the heatwave period caused by the stagnation event. Heatwaves can increase demand on the grid (Maia-Silva et al., 2020; Stone et al., 2023) and therefore increase NO_x emissions due to greater combustion of fossil fuels from power generation (Chen et al., 2015), which are not represented by the 2017 NEI. Given the potential low bias in modeled NO_x emissions, the high bias in modeled O_3 is even more perplexing, highlighting the complex chemistry involved in O_3 production. Importantly, the modeled MBE for O_3 is < 10 ppbv for central Missouri on June 29th, and Western New York on July 1st, locations that were upwind of WFM according to the HYSPLIT trajectories. This indicates that O_3 chemistry was well represented in the airmass that traveled to WFM.

Model predictions of PM 2.5 performed worse compared to O_3 with many linear correlation values exhibiting null or negative values and MBE exceeding 10 $\mu g\ m^{-3}$. Similar to O_3, model MBE was < 10 $\mu g\ m^{-3}$ for Missouri and much of Chicago on June 29th and Western New York on July 1st. Three air quality monitoring sites in New York, measuring O_3, PM2.5, and 2 meter temperature, were chosen for time-series evaluations of WRF-Chem, including Pinnacle State Park (PSP) in the Southern Tier of New York, Queens College in New York City, and measurements at the old ski lodge below the summit of WFM (Figure S9). More information about the data collected at these sites can be found in Brandt et al. (2016) and Ninneman et al. (2020), while Pearson correlation values and MBE statistics can be found in Figure S9. WFM tends to show the lowest linear correlation with observations. This is likely due to WRF-Chem underestimating the elevation of WFM (1483m) by over 700m, and therefore not properly accounting for the topography in the region (Figure S10). PSP shows the lowest MBE values with high correlation coefficients ($r > 0.7$) for O_3 and 2m temperature. Finally, Queens college saw the strongest correlation coefficients for O_3 and 2m temperature ($r > 0.85$), but exhibited large positive biases for O_3 and PM2.5. The causes behind these overpredictions remain unclear but are beyond the scope of this work."

We have also updated many sentences throughout the manuscript to be more quantitative when discussing model results.

Specific/Technical

Comments Line 7: remove "analysi".

This typo has been removed.

Line 89: "form" should be "from".

This typo has been removed.

Line 251: Remove space before period.

The space was removed.

Line 251: Add space after "trajectories".

A space was added after the word trajectories.

Line 258: Add period at end of sentence.

A period was added to the end of the sentence.

Line 331-332: Remove "(left)" and "(right)". This is appropriate for the figure caption and seems unnecessary in the main body text.

The words left and right were removed from the sentence.

Line 340: What do these hour numbers represent? Please clarify.

Figures 10 and 12 have been updated the x-axis to read "Local Time (EST)" for more clarity.

Figure 1: Please make font size larger. Even with zooming in, some words are too small to be legible.

An error in latex was causing the figure to appear smaller than the authors had originally intended. Figure 1 was enlarged to allow for easier reading.

Figure 5: The caption references a) and b) but the images are not labeled as such.

The figure caption was updated to remove the "a)" and "b)" references.

Figures 7 & 8: In the caption, "Plum" should be "Plume".

The word "plum" was changed to "plume" in each figure caption.

Figure 9: What does the blue solid line represent?

The blue line represents a fitted trend line using a generalized additive model (GAM) to estimate the average change of organic acid mixing ratios during the pollution event at WFM. A description of this trend line was added to the figure caption.

Supplement: Section S3 comes before S2? Line 152 references Section S3 for WRF Chem description but it is listed as S2 in the supplement. "Figures S8" under Section S6 should be Figure S7. In Section S7, "Figure S9" should be Figure S8.

The incorrect Section and Figure numbers have been corrected in the main text and supplemental material.