Response to Reviewer #2

We thank the reviewer for their comments on the manuscript. We have addressed these comments as described below. All reviewer comments are presented in italic font while the author responses are displayed in standard font. Specific text that was added to the updated manuscript is provided in blue text.

This is a review of the manuscript "Insights into the long-term (2005-2021) spatiotemporal evolution of summer ozone production sensitivity in the Northern Hemisphere derived with OMI" by Johnson et al. This paper is an important contribution to ongoing efforts to identify trends in surface air quality using satellite-based observations. This study investigates trends in column HCHO, tropospheric column NO2, and the HCHO/NO2 ratio (a.k.a. FNR) as observed by the OMI satellite. Overall, this is a good paper but needs some "polishing" and clarification.

In the discussion of OMI HCHO, it should be mentioned that Anderson et al., 2017 identified uncertainties in the use of the Tropical Western Pacific as a "clean" region when post-processing the OMI HCHO VCD.

The following sentence has been added to the OMI HCHO description section in the revised manuscript: "Model evaluation studies have shown that CTMs have errors and uncertainties in their predictions of HCHO in the clean regions of the Pacific Ocean which could also contribute to overall OMI HCHO bias/errors (Anderson et al., 2017).".

There are far more surface observations of NO2 than HCHO. Are the AQS NO2 and HCHO data co-located? If not, how is FNR calculated? Additionally, HCHO observations occur at 3, 8, 12 and 24 hr intervals and sometimes the HCHO data are only available every 6th or 12th day. Please provide more detail how these gaps are being handled and why only the 24hr data are used. Do you expect that afternoon OMI HCHO will be strongly correlated to 24hr avg surface observations? Is it appropriate to use 24hr average HCHO observations with 2hr avg, mid-afternoon NO2 data to calculate surface FNR?

The reviewer is correct that we only use AQS data that have co-located HCHO and NO_2 observations. This includes spatial and temporal co-locations. As the reviewer mentioned, sometimes AQS HCHO data is not available every day. AQS FNR values are only considered on the days for which HCHO measurements are available. Given the random distribution of the time of day in which AQS HCHO observations are taken, we have no other options other than using 24-hour AQS data. We don't calculate our own 24-hour averages as they are provided by the EPA already: https://aqs.epa.gov/aqsweb/airdata/download_files.html#Daily. These data files aggregate all HCHO observations taken throughout a single day and provide these daily 24-hour averaged HCHO values. We have added the following text to the revised manuscript in Sect. 2.2 to explain this: "Since there is insufficient hourly data for HCHO from the EPA AQS network, we use 24-hour average data for the HCHO evaluation which is provided by the EPA. AQS data for HCHO and NO_2 from each site are only used for days in which both species are measured.".

The reason we use normalized trends to compare OMI to AQS in situ FNR values is due to the issues the reviewer identifies. HCHO has diurnal variability which will not be replicated in 24-

hour averaged AQS data. HCHO data in surface in situ data suggests that HCHO values during the mid-day hours will be slightly higher than the daily average (Zhu et al., 2017). However, we look at normalized trends of interannual long-term variability which will not be impacted by whether we used mid-day or 24-hour average HCHO data in the AQS data.

Given there is no better option than to use 24-hour average AQS HCHO data, we use these values. However, for this study it is best to take advantage of the fact AQS data does provide hourly NO_2 data to compare mid-day NO_2 values to daily averaged HCHO information.

Are HCHO observations from the Interagency Monitoring of Protected Visual Environments (IMPROVE) network, the National Air Toxics Trends Stations (NATTS), and the Photochemical Assessment Monitoring Stations (PAMS) networks used for this this study?

The NATTS and PAMS measurement network information are included in the EPA AQS data applied in this study. We do not include measurements from the IMPROVE network since it only provides information on aerosol concentrations and speciation. Overall, for our study, we only apply AQS stations that have continuous HCHO and NO_2 observations throughout the study time period (2005 – 2021) which includes measurements from the NATTS and PAMS networks.

One issue for this study is the differentiation between urban and rural and how this component of the paper is incorporated into the study. The paper assumes that grid cells identified as "urban" are the same for every year but is this truly the case? For the 2005-2021 time series, is it possible that some "rural" areas become "urban"? If so, how will this impact the overall results.

The reviewer is correct that there has been urban expansion throughout the Northern Hemisphere over the last two decades. However, as mentioned in the response to the comment below, our urban category for this study includes CGLC-MODIS-LCZ land use categories 51 – 60, which includes a range of urban land use from sparely built to compact high-rise including the heavy industry category. CGLC-MODIS-LCZ land use categories between 56-59 represent suburban environments which are largely made up of green space with moderate to sparsely dispersed buildings. Therefore, conversion of suburban and sparsely built-up areas to more urbanized landscapes will already be included in our urban classification. The only thing not captured in our urban map will be transitions from completely vegetative lands to more built-up landscapes. This is expected to have relatively minimal impact on the results of this study. However, the following text has been added to the revised manuscript to clarify this issue: "The CGLC-MODIS-LCZ urban and rural maps derived for this study are static and will not capture urban expansion which has occurred over the last two decades. However, since our urban classification includes both urban and suburban landscapes (including sparsely built-up areas), the transition from suburban to urban landscapes will already be included in our urban map. The only thing not captured would be the transition from completely vegetative areas to more built-up landscapes which are expected to have a minor impact on the results of this study.".

Suburban and urban are lumped together as "rural". Is this appropriate? In some areas, suburbs have large populations and/or are along major interstates and suffer from significant, local pollution emissions. Do the results reported here change if only truly rural areas are considered?

I would expect suburban areas to be influenced by both urban and/or rural depending on meteorology. I think additional discussion is warranted.

In the original manuscript we provide the following description: "Urban classification is defined by the CGLC-MODIS-LCZ land use categories 51 – 60, which includes a range of urban land use from sparely built to compact high-rise including the heavy industry category". CGLC-MODIS-LCZ land use categories between 56-59 represent suburban environments which are largely made up of green space with moderate to sparsely dispersed buildings. Therefore, the "urban" category used in this study includes urban and suburban landscapes. We feel that including suburban areas in the "urban" classification is appropriate for the reasons listed by the reviewer. For clarification the following sentence has been added to the manuscript: "Urban classification is defined by the CGLC-MODIS-LCZ land use categories 51 – 60, which includes a range of urban land use from sparely built to compact high-rise including the heavy industry category. These land use categories capture both urban and suburban landscapes.".

Table 1 is confusing. Please be more specific as to what is being presented. Are the "obs." referring to OMI or AQS sites? What is the "model"? Perhaps I missed it but what is the model referring to? CEDS? Also, which of these statistics are actually significant? A correlation coefficient (R) of -0.27 is an R^2 of 0.07, which is quite small.

The title of Table 1 has been updated to clarify the AQS data is referenced as "Obs.": "Table 1. Statistics of the correlation of OMI and AQS ("Obs.") normalized trends for HCHO, NO_2, and FNRs for major cities in the US and the average of all cities in the US (USA urban areas) between 2005-2019. Slopes of the trends for each species (units of yr^{-1}) are also provided. The values in italic font are linear regression slopes which are statistically significant at a 95% confidence level ($p \leq 0.05$).". Furthermore, "Model" has been replaced with "OMI" to clarify this is the satellite observations. This was a typo in the original manuscript.

Table 1 has been updated so that linear regression slopes for AQS observations and OMI retrievals are highlighted for statistical significance values of $p \leq 0.05$ (95% confidence level). The table caption has been updated with the following sentence to highlight this as shown above. Some additional text has been added in Sect. 3.2 in the revised manuscript to highlight the fact that all AQS and OMI NO_2 and FNR linear regression trends were statistically significant at a 95% confidence level.

Figure 4: The y-axis changes for some of the panels. Is it possible to have a uniform Y-axis throughout?

This figure has been corrected as suggested by the reviewer.

Figure 7&8: Please choose more distinct colors to make it easier to discern between the 3 years ranges or years. The bars are very narrow and it's difficult to clearly see the difference between orange and gold.

We have altered the color scheme of the bars in Fig. 7 and 8 as suggested by the reviewer.

Figure S3: The city names on some of the panels overlaps with the "10^15" label for the Y-axis. You can overcome this by removing "10^15" from each panel and simply including this in the Y-axis label, i.e. 10^15 Molecules/cm^2"

This has been corrected in the revised supplemental material.

References

Anderson, D. C., Nicely, J. M., Wolfe, G. M., Hanisco, T. F., Salawitch, R. J., Canty, T. P., Dickerson, R. R., Apel, E. C., Baidar, S., Bannan, T. J., Blake, N. J., Chen, D., Dix, B., Fernandez, R. P., Hall, S. R., Hornbrook, R. S., Gregory Huey, L., Josse, B., Jöckel, P., Kinnison, D. E., Koenig, T. K., Le Breton, M., Marécal, V., Morgenstern, O., Oman, L. D., Pan, L. L., Percival, C., Plummer, D., Revell, L. E., Rozanov, E., Saiz-Lopez, A., Stenke, A., Sudo, K., Tilmes, S., Ullmann, K., Volkamer, R., Weinheimer, A. J., and Zeng, G.: Formaldehyde in the Tropical Western Pacific: Chemical Sources and Sinks, Convective Transport, and Representation in CAM-Chem and the CCMI Models, J. Geophys. Res.-Atmos., 122, 11201–11226, https://doi.org/10.1002/2016JD026121, 2017.

Zhu, L., Jacob, D. J., Keutsch, F. N., Mickley, L. J., Scheffe, R., Strum, M., González Abad, G., Chance, K., Yang, K., Rappenglück, B., Millet, D. B., Baasandorj, M., Jaeglé, L., and Shah, V.: Formaldehyde (HCHO) as a hazardous air pollutant: mapping surface air concentrations from satellite and inferring cancer risks in the United States, Environ. Sci. Technol., 51, 5650–5657, https://doi.org/10.1021/acs.est.7b01356, 2017.