This paper promises to “critically assess the efficacy of these [NO\textsubscript{x}/CO\textsubscript{2}, CO/CO\textsubscript{2}] ratios to isolate the combustion contribution of CO\textsubscript{2}.” This is an interesting and timely goal. After reading the paper, which contains high-quality figures and is written fluently, I can only conclude that paper mostly fails to reach this goal.

Methodologically, the paper lacks proper hypotheses about what to expect, and has a substantial number of flaws. The GEOS-Chem model does a poor job to simulated NO\textsubscript{2} columns over major CO\textsubscript{2} emitting areas within Europe. This makes the analysis of the paper rather speculative. I attach an annotated pdf with comments, and below I summarize the main points that need to be addressed before it can be considered for publication.

- Formaldehyde and December introduced but not presented.

In the introduction, the authors introduce formaldehyde as a species that will be studied: “we explore the agreement between model and observed ratios of NO\textsubscript{2}, CO, and HCHO with CO\textsubscript{2}.”

They write: “Formaldehyde (HCHO) is another proxy for incomplete combustion (e.g., Fu et al. (2007a); Gonzi et al. (2011))”. However, later HCHO is discarded: “Given the limited use of HCHO as a tracer of combustion we do not pursue this tracer any further.” Why include this tracer anyhow? I understand that the authors worked on HCHO before (they refer to their papers), but the added value of showing HCHO in this paper is limited, because this paper links emission ratios to modeled atmospheric ratios. Since HCHO does not have substantial emissions, I would recommend the authors to leave this tracer out altogether. As a side note, one could think of a way to use HCHO to quantify the interfering contribution of VOCs to the CO budget (an issue that is left largely open). The authors write on page 9: “Using GEOS-Chem, we find that the secondary source is
typically 10–20% of the total CO source in winter months but in July can be as much as 75% of the total CO source over Europe. This secondary source will therefore need to be considered if CO is to be used to isolate combustion CO\(_2\).” One reason to leave HCHO in, would be an analysis of how HCHO could be used to quantify the indirect CO source.

On a similar note, satellite data are not useful in December 2018, because there are insufficient valid OCO-2 observations. The presentation of the December emissions only does not supply much added value. In fact, the analysis is based on only one month of model simulation and satellite data, which is rather limited.

- CO\(_2\) modelling.

Concerning the modelling of CO\(_2\), the description is very brief. It seems that the authors do not include the exchange with the biosphere (or oceans). Yet, in figure 5 they compare OCO-2 with the model. Given the large role of the biosphere in July, I would argue that this analysis is therefore flawed.

- Sharp gradients in the emission gradients.

The description of the emission ratios applied in the model reveals sharp boundaries between the countries (also in Super et al. 2020). This can partly be understood (“The differences between countries for the two months reflect the relative importance of individual sectors”), but the authors should at least mention that there are still significant uncertainties in reporting of national scale emissions, their allocation, and downscaling. It is hard to believe that emission ratios suddenly change when you cross a border.

- Strategy and results.

The authors mention a high spatial correlation between GEOS-Chem and TROPOMI (R=0.60 for NO\(_2\) and R=0.82 for CO). However, this spatial correlation in \(\mu\)mole/m\(^2\) includes a spurious correlation driven by orography in the domain. I suggest to recalculate the correlation in mixing ratio units.

The follow-up question is then of course: how good is this? Visually, it does not look great. The fact that important hotspots, like Madrid, Paris, and London, are misrepresented by the model does not give large credibility to the simulations. Has this model version been validated against surface observations? How can I be sure that the chemistry is not totally
flawed over urban environments?

Simply giving reasons is not enough ("...likely reflect over-reporting of NO\textsubscript{x} emissions from rural areas of France, Germany, Poland and other eastern European countries, i.e., errors in emission inventory, temporal profile and errors in vertical mixing and lifetime of NO\textsubscript{x} against chemical oxidation."). Which is due to some combination of underestimating emissions from these large urban sources, errors in the model description of NO\textsubscript{x} photochemistry, and the low sensitivity of averaging kernels to lower levels of the atmosphere.

Basically, we are dealing with largely unvalidated model results.

Nevertheless, the authors proceed ("....provide us with some confidence in our ability to use enhancement ratios of column CO and NO\textsubscript{2}."). After some manipulation (subtracting background from model and satellite data) they present in Figure 6 regional enhancements on the national scale. The good correlation they find for CO\textsubscript{2} and CO on country level is reassuring, although it remains unclear why they moved to the country scale.

However, for NO\textsubscript{2} the modeled columns are too small in the Netherlands, Belgium, and Germany, and too high for Ireland. This calls at least for some interpretation. Likely, the smearing of the high NO\textsubscript{x} emissions leads to a too-high OH regime. Also, one should question the use of the standard TROPOMI averaging kernel (from the coarse-scale TM5 model) for the higher-resolution simulations (see abundant literature on this subject).

The authors now take the next step in figure 7, in which they show the ratio of the NO\textsubscript{2} and CO enhancement relative to the CO\textsubscript{2} enhancements on the grid-scale (back to the grid scale?). My problem with this procedure is that the spatial variations in NO\textsubscript{2} (typically 25%, from figure 6) and CO (typically 15%) are dominating the CO\textsubscript{2} signal (variations < 10%). A more serious problem is that the errors associated with the modeled NO\textsubscript{2} variations are larger than the signal in CO\textsubscript{2}, making the analysis that follows flawed in my opinion.

In the next step, the authors proceed one step further, by correlating (figure 8a) the combustion ratios NO\textsubscript{x}/CO\textsubscript{2} with CO/CO\textsubscript{2}. There is, however, not a priori hypothesis why there should be a correlation. Do countries with a large proportion of sectors with high CO/CO\textsubscript{2} emission ratios, also have high temperature combustion (with high NO\textsubscript{x}/CO\textsubscript{2} ratios? Apparently not, because there is too much scatter to draw a correlation line. Figure 8a does not bring relevant information here, and would belong to section 3.1. Concerning the unit of the slope: I think it is not unitless (kg NO\textsubscript{x}/kg CO, and a factor 10 somewhere).

Figure 8b does bring interesting information. But, if you divide both the x-axis and the y-axis by the same number (ΔCO\textsubscript{2}), would it not be more straightforward to simply study the ΔNO\textsubscript{x}–ΔCO relation? Earlier identified model flaws in correctly modelling TROPOMI NO\textsubscript{2} mostly explain the difference in slope between model and satellites. The added value
of dividing by the CO\textsubscript{2} enhancement is not clear, nor quantified. The authors write: “This suggests both observed and model columns reproduce the positive relationship based on the inventory estimates.”, and basically confirm that CO and NO\textsubscript{x} emissions are correlated both in the inventory and in the modeled columns (since CO\textsubscript{2} divides out).

The final figure (figure 9) correlates the country scale emission ratios to model and satellite ratios. Interestingly, for NO\textsubscript{2}, a negative correlation is observed. This would imply that countries with high NO\textsubscript{x} emissions (compared to CO\textsubscript{2}) show smaller NO\textsubscript{2} columns (compared to CO\textsubscript{2}), which seems rather surprising, and warrants further interpretation. The authors interpret: “the negative relationships between the inventory-based and atmospheric-based ratios reflect a strong non-linearity between NO\textsubscript{x} emissions and NO\textsubscript{2} concentration.” In trying to understand this issue better, I went back to figure 4, noting from figure 9 that Norway (2.7x10\textsuperscript{-3} mole NO\textsubscript{x}/mole CO\textsubscript{2}), Spain (2.4x10\textsuperscript{-3}), and Portugal (2.8x10\textsuperscript{-3} mole NO\textsubscript{x}/mole CO\textsubscript{2}) show the largest emission ratios, and that Germany (1.3x10\textsuperscript{-3} mole NO\textsubscript{x}/mole CO\textsubscript{2}) has the lowest ratio in figure 9. Indeed, in figure 4 Portugal has the highest emission ratio, but values and order are distinctly different from figure 9 (Poland is for some reason not shown). Do I interpret the numbers wrong? Anyhow, this issue requires some further scrutiny.

In conclusion, although the subject is interesting and timely, the paper needs substantial revision, scrutiny, and would benefit from a complete redesign. A thorough model validation, including comparisons to surface observations would be needed to gain trust in the model. The severe underestimate of NO\textsubscript{2} over major cities in Europe precludes the analysis the authors nevertheless present. Also, I would recommend the authors to improve the comparison to TROPOMI NO\textsubscript{2}, by recalculating the averaging kernel based on the high-resolution NO\textsubscript{2} distribution (see literature).

Please also note the supplement to this comment: https://acp.copernicus.org/preprints/acp-2021-816/acp-2021-816-RC1-supplement.pdf