Comment on acp-2021-121
Anonymous Referee #2

Referee comment on "Unraveling Pathways of Elevated Ozone Induced by the 2020 Lockdown in Europe by an Observationally Constrained Regional Model: Non-Linear Joint Inversion of NOx and VOC Emissions using TROPOMI" by Amir H. Sourí et al., Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2021-121-RC3, 2021

Major comments

This manuscript presents an inverse modeling study of NOx and VOC emissions over Europe in the spring of 2019 and 2020, based on TROPOMI NO2 and HCHO data. The focus is on the differences between the two years and on the detection of Covid-related effects. In agreement with previous studies, large NOx emission decreases are derived over most European countries in April 2020. In March and May, however, the picture is less clear, with some regions (e.g. large parts of UK and Germany in March) showing large emission increases in 2020 (Figure 4). Those changes are unrealistic and are contradicted by comparisons with surface NO2 measurements (Figure 7). The authors present the disparity between regions as a consequence of the different timing of Covid-lockdown measures over different regions of Europe, but the discussion is poor and does not present concrete arguments for the inferred patterns. Those patterns are probably related to the inability of the model to match the observed NO2 column distribution, in particular (but not only) over N-W Europe (see Figure S3). The prior model strongly overestimates NO2 over northern Germany and strongly underestimates NO2 in southern Germany and in many other regions. It would be most enlightening to examine the top-down emission increments for each year and month (and not just the differences between the two years). I suspect there will be huge disparities within several countries, especially Germany. It is highly unlikely that bottom-up emission inventory could perform so badly in terms of spatial distributions.

The model compares also very poorly against TROPOMI HCHO, although this is partly due to issues with the observations. As detailed further below, the high HCHO values in April (also May) in Scandinavia and Russia might be artefacts as evidenced by comparisons of TROPOMI with FTIR data (Vigouroux et al. 2020). The authors apply a crude "bias correction" to the data (decrease by 25% the values below 2.5E15 cm-2, increase by 30% the values above 7.5E15 cm-2) but it is inappropriate as it probably increases values which are already too high (e.g. in April over parts of Russia and Scandinavia) and it leads to different corrections being applied in 2019 and 2020, thereby creating artificial patterns in the differences between the two years. Given the magnitude of the biases between the model and the satellite, the inferred top-down emission differences (Figure 5) have no credibility at all. Top-down anthropogenic VOC emissions in April appear to be much higher than in March, without justification. In the abstract and conclusions, much case is
made of the small VOC emission decrease in March 2020, compared to March 2019, as if it
could be related to Covid. This is highly misleading since the averaging kernel (AK, Figure
5) is lowest (close to zero) in March, i.e. the decrease is not well constrained by the data.
The authors emphasize very much the importance of the NOx inversion on the results for
VOCs in March, but their arguments (that VOC emissions in Spain and Italy are decreased
only to compensate for the stronger VOC oxidation above cities due to NOx decreases) do
not stand scrutiny. As shown by Figure S5, the inversion does increase HCHO columns
over Spain and Italy in 2019, bringing the model (a little bit) closer to the data. For some
unknown reason, this HCHO increase does not happen in 2020 despite similarly high
TROPOMI columns. This explains the March patterns of Figure 5. There is an influence of
the NOx optimization on the VOC results, but probably not the onedescribed by the paper.

The paper is too long and, at many instances, not clearly written. I have provided a
number of suggestions for improvement, but I encourage the authors to make a general
effort towards more clarity. Many sentences and entire paragraphs are given which do not
add much to the discussion. For example, Figure 6 is very long to describe, but I am not
sure whether it really helps to interpret the results. If it does, please make it more clear
and remove unnecessary parts. On the other hand, much information about the model
and methodology is incomplete or clearly wrong (e.g. the adopted errors for TROPOMI
HCHO).

In conclusion, I do not recommend the paper for publication (in its present form) since its
conclusions are not well supported by analysis of the data. I recommend to scale down the
ambition of the paper. The HCHO (and AOD) data do not seem to help constraining the
emissions. The NOx part could be interesting if presented honestly with its caveats.
Sensitivity inversions would help to appreciate the uncertainties and robustness of the
conclusions.

Minor comments

Abstract: very long, should be shortened

1 122-124 "Since vertical column densities (...° depend on assumed gas profile shape (...),
we recalculate those shape factors using profiles from our (...) model": the air mass
factors being a complex function of profile shapes, cloudiness, albedo, etc., more details
are needed to describe how the profile shapes are taken into account.

1 156 Why the RMSE? Do you mean the assumed uncertainty on the NO2 columns from
TROPOMI?

I 156 The values of 1.1E15 and 3.5E15 molec cm-2 seem arbitrary. Please provide better
explanation of how those were derived (as they play an important role in the emission
inversion)

1 176-178 "Vigouroux et al (...) majorly located over pristine areas and 9 MAX-DOAS
stations...": wrong. The paper concerns FTIR stations, not MAX-DOAS. Furthermore,
many of the FTIR sites are in cities (e.g. Paris, Bremen, Mexico city, etc.). Please check
the references you cite.

I 181-183 "The agreement between MAX-DOAS...": again, this paper concerns FTIR data
only. Please provide the correct references to your statements.

I 184-186 Please provide the precise procedure used for deriving those numbers.

I 186-187 A value of 4% seems extremely low and unlikely given the large biases and
scatter of the FTIR-TROPOMI comparisons.
The motivation for using MODIS AOD is not made clear. Clarify. If it does not bring anything, why this complication?

The assumption that the interferences are similar in 2019 and 2020 due to low photochemistry is crude. In Lamsal et al 2008, the correction factors in spring over the U.S. range typically between 0.4 and 0.7. Since CMAQ calculates the interfering species (PAN etc.), why don't you apply the correction proposed by Lamsal et al.? It is a rough correction but it would be better than no correction at all.

Do you use gridded maps of the emission factors or PFT distributions in conjunction with the emission factors from Table 2 in Guenther et al 2012?

Does the model include soil and lightning NO emissions? The use of fertilizers could be a significant source in spring.

Is the diurnal cycle of anthropogenic emissions taken into account? Regarding the biogenic emissions from MEGAN, are diurnal and day-to-day variations included? What VOC species are emitted (besides isoprene)?

Is a collection of monthly NO2 columns or daily NO2 columns? Is the model sampled at the satellite overpass time? Specify the temporal tolerance window.

Why three times? How do know whether this is sufficient?

The rationale for this assimilation of MODIS AOD is not clear.

The faster vertical mixing should generally lead to higher NO2 columns due to the higher sensitivity of TROPOMI to NO2 at higher altitudes. Stronger advection does not change much when averaged over a sufficiently large area. Clearly, the increased photochemical activity is by far the main reason for lower NO2 columns in later months.

"we see negligible reductions...": there is no reduction at all. There is a significant increase in these regions (boxes B, C and D). Rephrase.

"northern Germany is associated with less populated areas": quite an extraordinary statement. Please look at population density maps. Please focus on relevant information, e.g. the timing of the lockdown, besides meteorological variability. When did lockdown measures take effect in Germany, France, Italy, etc.?

A detection limit of 7E15 cm-2 is not "very low", since it is higher than the TROPOMI columns at most locations in March-May (Figure S5)

The reference Karlsson et al. 2013 does not inform on the occurrence of biomass burning in 2019

Over St Petersburg, the FTIR HCHO column in April 2019 is about 4.2E15 cm-2 (Vigouroux et al. 2020), a factor of 1.6 below the TROPOMI column. A similar overestimation is found at Kiruna. In May, the discrepancy is even higher at Scandinavian sites. Clearly, TROPOMI data over Northern Europe and Russia in spring need to be considered with extreme caution. The "dipole anomaly" (line 319) might very well be an artefact (at least quantitatively)

"the fact that isoprene reactivity significantly increases by rising temperature": the OH-rate constant actually decreases at higher temperatures. The chemical lifetime of isoprene is always short enough that it is oxidized close to the emission area. Nevertheless, there is a longer delay in winter/spring before oxidation products like MACR
and MVK get oxidized and form HCHO. But this should not play a significant role compared with the temperature-dependence of biogenic emissions. Note furthermore that over Russia and Scandinavia, where coniferous trees are dominant, monoterpenes (not isoprene) might be the main biogenic precursors of HCHO. Are those emissions (and their subsequent chemical oxidation) considered in the model? If not, what could be the consequence of their omission?

Section 3.2 Before discussing the top-down emissions, the paper should discuss the performance of the a priori model against satellite data. I find striking that the model fails at reproducing many prominent features of NO2 column distributions. Why is CMAQ NO2 so high along the coasts of Germany and Holland whereas it is notably too low e.g. over southern Germany? Over Ukraine and other regions, the model is too low by a very large factor (>4). The paper should show not just the top-down emissions but also the emission increments and discuss whether those increments have any plausibility. I have serious doubts on that matter. The a priori emission distribution (from CEDS) might have some uncertainty but cannot be completely wrong.

I 328 "elsewhere": elsewhere in the paper or in a further study? I would guess that these aerosol changes have only limited impacts on NO2 and HCHO. If so, the impact of AOD assimilation should be either briefly mentioned or dropped entirely from the paper. If not, it would be interesting to discuss more in detail.

I 335-336 "large reduction (...) in the bias associated with simulated surface NO2": why not show this, e.g. in the Supplement?

I 338-339 "the discrepancies between the simulated tropospheric NO2 columns versus TROPOMI are largely mitigated by the inversion": only in region with highest emissions, not at all elsewhere.

I 342 "because of the consideration of observational errors": but the choice of NO2 column errors was pretty arbitrary (as far as I understand). It could be useful to show inversion results adopting alternative choices of those errors and other setup parameters.

I 343 During summer and even in spring (at least in southern Europe), the feedback would be: if NOx increase, then OH increase, then the NOx lifetime decreases, implying a larger NOx emission increase is needed to match the NO2 enhancement. Therefore, it does not seem obvious that chemical feedbacks would decrease the magnitude of the anomalies. Please clarify, or drop the mention of chemical feedback.

I 344 "some doubt the practicality of direct mass balance methods": at least, such methods provide a direct answer independent on assumptions regarding uncertainties.

Table 2: The absolute differences of top-down NOx emissions are not really useful and could be dropped.

I 358 As for NOx, a discussion of the model performance is needed for HCHO, before discussing the inversion. In addition, the a priori VOC emissions should be shown for the 3 months. Generally speaking, there is a huge underestimation of the model against TROPOMI HCHO (Fig S5). That might be partly due to biases in the data (see above regarding FTIR vs TROPOMI comparisons) but should clearly be mentioned. My guess is that the model would underestimate the FTIR HCHO columns at sites like Paris and Bremen. In any case, the large uncertainties in TROPOMI HCHO make the inversion results unreliable (except maybe at low latitudes in May). The differences "Lockdown-Baseline" (Figure 5) are even more uncertain. I think they should not be shown at all as they might mislead the reader.
This tendency, which is striking, mainly stems from the indirect impacts of the reduced NOx emissions on HCHO: the reasoning is overly simplistic. E.g. over Spain, the largest change is not seen over Madrid but in an area to the west of the city. Over Italy as well, the changes are spread over wide areas. Furthermore, in April the VOC emissions are found to increase quite a lot over cities like Paris, Rome, Milano, etc. Obviously the patterns are primarily dictated by the large differences between TROPOMI HCHO and the model, despite the large HCHO uncertainty adopted in the inversion.

An additional worry concerns the seasonality of top-down emissions in 2019. According to Fig 5, VOC emissions in April are considerably higher than in March over most countries. The retrieved emission patterns indicate primarily anthropogenic emissions. How can this be justified?

Section 3.3.1 could be shortened. Get to the point!

In March, surface NO2 is higher in 2020 than in 2019 according to the model (not the data), consistent with the column changes (Fig 4). This confirms the suspicion that the NOx emission changes in these regions are unreliable due to large model errors.

Figure 8: is this given for 2019 or 2020? I suggest providing both years, but in the Supplement instead of the main article.

"large spatial and temporal variability associated with the reduction in NOx was evident as each country might have different level and timeline of restrictions": however, the discussion of this aspect is poor in the paper. "emissions decreased in April rather than March in some portions of UK, northern Germany...": do you really mean that the UK and Germany both showed significant regional differences in the lockdown measures? I doubt very much that it was the case, but if it is, it should be discussed and better justified.

"we showed that anthropogenic VOC emissions over Paris (...) decreased in March (...) achievable through jointly using NO2 and HCHO observations" as noted above, this is very doubtful. You have not made your case that the VOC emission changes are due to NOx emission changes. For that, you should realize a separate inversion using only TROPOMI HCHO and compare with the standard inversion.

Technical/language comments

"estimate of the NO2 reduction is underestimated": rephrase

"a picture that correlates with the TROPOMI etc.": unclear

"TROPOMI HCHO sets an upper limit for HCHO changes such that the chemical feedback (...) reveals a non-negligible decline...": unclear. That a feedback reveals a decline doesn't make sense.

"Results of integrated process rates of MDA8 surface ozone": unclear

"capture the essential character changes..." essential in what sense? Unclear.

"has exponentially become more polluted during previous decades": wrong over Europe

"impulsive and sweeping": not clear what is meant

"particulate matter" (drop the s)
Why the indentation?

low spatial resolution (remove hyphen)

"while considering" unclear

The sentence should make more clear that "clear" is <6x10^{15} molec/cm^2 and polluted is above that level. Use the proper symbol for >=

"Those biases oscillates around 8x10^{15} molec/cm^2": completely unclear.

"majorly" -> mainly

The sentence "We nudge moisture (...) data used only outside of the PBL layer" is a bit ambiguous, please rephrase

"PBL layer" is redundant

the correspondence is good, not striking.

are assumed diagonal

"unintended" is weird. NO2 columns have no intention. Rephrase.

"are first the free-tropospheric region complication": what does this mean? Not clear at all.

"a barrier to obtaining high amount of information from the sensor..." unclear, rephrase.

"suggests an abrupt hiatus in the ongoing reduced NOx emissions": unclear

Why "potential"?

"leading to striking HCHO column patterns with large variations" does not tell anything, please remove.

"higher chance": is it really a matter of chance? Rephrase.

"looking at": rephrase

"are below the detection limit (...) to relate them to the lockdown...": lousy wording, please rephrase (e.g. remove the last part of sentence"

"nonetheless TROPOMI sets an upper limit of these changes": not useful

"we surprisingly observe": weird wording

"ignoring spatial representivity function to directly compare point measurements...": unclear, rephrase

"then are then"

"heterogenicity" --> "heterogeneity"
The surface measurements reinforce the less pronounced reduction in NO2 in northern Germany and UK: unclear.

The sentence "This tendency potentially is driven (...) has drawn much attention" is grammatically incorrect.

"The challenge is to simulate a model": unclear

"essential character": unclear, rephrase

"namely as": delete "as"

In Equation (6), the rate constant should be k(O3+NO), not k(OH+NO2+M)

Remove comma at end of sentence

"explain" --> "describe"

Table 1 hyphen in MAX-DOAS

Table 2: too many significant digits are given.