Anonymous referee #1

Anonymous during peer-review: Yes No

Anonymous in acknowledgements of published article: Yes **No**

Checklist for reviewers

1) Scientific significance Does the manuscript represent a substantial contribution to scientific progress within the scope of this journal (substantial new concepts, ideas, methods, or data)?	Outstanding **Excellent** Good Fair Low
2) Scientific quality Are the scientific approach and applied methods valid? Are the results discussed in an appropriate and balanced way (consideration of related work, including appropriate references)?	Outstanding **Excellent** Good Fair Low
3) Presentation quality Are the scientific results and conclusions presented in a clear, concise, and well structured way (number and quality of figures/tables, appropriate use of English language)?	Outstanding **Excellent** Good Fair Low

For final publication, the manuscript should be

accepted as is

accepted subject to **technical corrections**

accepted subject to **minor revisions**

reconsidered after **major revisions**

rejected

Were a revised manuscript to be sent for another round of reviews:

I would be willing to review the revised manuscript.

I would not be willing to review the revised manuscript.

Suggestions for revision or reasons for rejection

(visible to the public if the article is accepted and published)

Paper is acceptable as is for publication.

Response: We appreciate reviewer's help to improve our manuscript.

Anonymous referee #2

Anonymous during peer-review: **Yes** No

Anonymous in acknowledgements of published article: **Yes** No

Checklist for reviewers

1) Scientific significance Does the manuscript represent a substantial contribution to scientific progress within the scope of this journal (substantial new concepts, ideas, methods, or data)?	Outstanding Excellent **Good** Fair Low
2) Scientific quality Are the scientific approach and applied methods valid? Are the results discussed in an appropriate and balanced way (consideration of related work, including appropriate references)?	Outstanding Excellent **Good** Fair Low
3) Presentation quality Are the scientific results and conclusions presented in a clear, concise, and well structured way (number and quality of figures/tables, appropriate use of English language)?	Outstanding Excellent Good **Fair** Low

For final publication, the manuscript should be

accepted as is

accepted subject to **technical corrections**

accepted subject to minor revisions

reconsidered after **major revisions**

rejected

Were a revised manuscript to be sent for another round of reviews:

I would be willing to review the revised manuscript.

I would not be willing to review the revised manuscript.

Suggestions for revision or reasons for rejection

First of all I thank the authors for their detailed replies to my comments and changes made to the manuscript. These have, to my opinion, improved the manuscript considerably, in particular the rewritten sections 2.4, 3.3 and in particular section 4.2 address several of my comments. The addition of the uncertainty estimate from the emission ensemble spread is also of interest. But I still have a couple of remaining and unanswered comments which I would suggest the authors take into account before the paper is published.
Response: We have tried to answer all comments posed from the reviewer step-by-step in the manuscript with Track Changes enabled.

The Tichy 2022 preprint on the COVID-19 impact on ammonia nicely shows the importance of SO2 and NO2 in determining the lifetime of ammonia. The COVID-19 reductions in NOx and SO2 emissions have a major impact on the lifetime, as mentioned in the last lines of section 3.1 (which is neglected in the current paper). And lifetime has a major impact on the emission estimate. For trend estimates it is clear that a wrong representation of trends in NOx and SO2 will lead to systematic biases in the estimated trends. So it is very important to demonstrate that the modelling system correctly represents the trends in NO2 and SO2. As mentioned by the authors use is made of state-of-the-art inventories, which provides some trust in the trends presented. But the realism of the lifetime estimate could be demonstrated by comparing modelled NO2/SO2 time series (trends) with surface observations of these two species, and by presenting this as for instance an extra figure in the supplement. In the paper NOx and SO2 are only mentioned in the disclaimer on COVID-19, but are not discussed at all in the results/discussion parts of the paper. My major comment was about this lack of discussion in the paper about this key factor in the trend estimate. Please add a discussion and possibly observational evidence that the NOx/SO2 levels and trends are reasonalbly well modelled in the system.
Response: We understand the concerns reported by the reviewer and we have now added an extended validation of NO2 and SO2 as requested. We have added 2 example plots validating ground measurements of SO2 and NO2 (respectively) against modelled concentrations in random stations and scatterplots of the full dataset used for the validation (Suppl. Figure S 8 and S 9). We also discuss the model validation with regard to NO2/SO2 concentrations in the paragraphs starting at P.15-L.479.

My major comment "The assumptions/modeling of errors of the satellite data (including filtering), in the method (model uncertainty: chemistry, transport) and a-priori emissions are not described" has not been fully addressed.
Response: We have tried to address the comments from the reviewer in the points below.
However, please note that uncertainty of the prior emissions has not been calculated in detail. Klimont et al. (2017) (http://www.atmos-chem-phys-discuss.net/acp-2016-880/) reported “We have not performed a formal uncertainty analysis for emission estimates in this study, but results of analysis from other studies are helpful and indicative of the expected un- certainties for various species and regions. For example, the global BC and OC inventory developed by Bond et al. (2004) included an uncertainty analysis of total emissions providing regional “low-high” estimates for 1996. For BC emissions from

anthropogenic sources, the range was 3.1–10 Tg yr−1 (−30 to +120 %) and for OC 5.1–14 Tg yr−1 (−40 to +130 %)."

1. Satellite data: The satellite data is "interpolated" to a 0.5 by 0.5 degree grid, reducing the number of observations. Details on how this is done are missing. A linear interpolation is mentioned. But are the kernels and covariance matrices interpolated in the same way? Please explain how this is done. Is the error (covariance) of a weighted mean concentration of various observations equal to the weighted mean of the covariances? In other words: are nearby CrIS observations correlated or uncorrelated? Is the full error covariance of CrIS taken into account (errors are strongly correlated)? Please specify how the satellite observational operator has been generated "in a robust manner."

Response: We have changed the term "interpolation" to "gridding", because it represents better what we did and we further explain in detail the procedure (Track Changes P.5-L.133-140). Furthermore, we give 8 supplementary figures (Supplementary Figure S 1), corresponding to 1st September of each of the 2013-2020 years of our study period, showing the quality of the gridding with respect to raw kernel data, gridded on 0.5 degrees over Europe, and the calculated standard deviation of the values falling within each 0.5 degree grid-cell. The calculated low standard deviation as compared to the gridded AK values show that the raw kernel values were very similar in each grid-cell causing small bias.

2. Modelling uncertainties: Are the grid-cell specific SRMs assumed to be free of error? Section 2.4 provides technical details on the implementation but does not discuss input assumptions for the model-related uncertainties. R in eq.5 is modelled with an omega factor and I_p. What is this I_p? Is it a diagonal unity matrix (are correlations between the vertical levels in CrIS accounted for)?

Response: I_p is the identity matrix with ones on its diagonal and zeros otherwise, while factor omega is estimated by the model within the calculation procedure. This is now clarified in the revised version of the manuscript (please see manuscript with Track Changes at P.8 – L.220-236).

3. Emissions: A cost function (equation 3) normally contains a term reflecting the a-priori uncertainty (of the emissions). This seems to be missing here. Are emission uncertainties taken into account in the optimal estimation?

Response: The optimization formulation in Eq. (3) and its restriction to the surroundings in Eq. (4) is formulated without uncertainty terms. However, a probabilistic model is employed exactly for the introduction of uncertainties into the inference. Here, we refer to Tichý et al. (2020) paper, where equivalence between classical optimization formulation and its probabilistic counterpart is shown, with the benefit of a probabilistic model, which can adaptively estimate other parameters of the model such as the covariance matrices, namely omega, L and V. Note also that the structure of the covariance matrix, e.g., in Eq. (7) is our prior assumption, however, its posterior structure can differ from the prior structure significantly. Typically, the posterior structure is a full matrix, see Appendix B in Tichý et al. (2020) for more details on posterior parameters calculation.

In the reply the authors mention "A diurnal cycle is neither assumed in the Chemistry Transport Model, nor exists in the satellite observations from CrIS". Please mention this

point explicitly in the discussion section of the paper. The monthly-mean emissions probably need to be interpreted as montly-mean emissions at the satellite overpass time, which will differ from the 24h daily mean emission.
Response: We have added this statement in Track Changes P.9-L.274-276. This is a common problem in many similar studies that involve satellite observations. We hope this specific comment will not cause further confusion to readers.

I did not understand equation 9. Please define the quantities in this equation (sigma_location). Does the square root cover the whole expression? I would expect a 1/n normalisation before the summation.
Response: We thank the reviewer for this question/comment. We reformulated equation (9) by replacing the $\Sigma_{location}$ to the $\sigma_{location}$ to avoid confusion with the sum symbol used in the right side of the equation (Track Changes P.13-L388).
Please note also that there is no normalization before the summation. This arises as a property of the sum of Gaussian distributions. For Gaussian distributed independent variables X_1, X_2, …, X_n with means mu_1, mu_2, …, mu_n and variances sigma_1^2, sigma_2^2, …, sigma_n^2, their sum follows the normal distribution:
sum_{i=1}^n X_i = N(sum_{i=1}^n mu_i, sum_{i=1}^n sigma_i^2)
When reporting the standard deviation rather than variance, the standard deviation is square root of the variance, hence sqrt(sum_{i=1}^n sigma_i^2), which is the core of equation (9).

Sec 2.1, line 157 (v2 manuscript) mentions "The individual profile retrieval levels show an estimated random measurement error of 10–30 %, with total random errors estimates increasing to 60 to 100%" What is the difference between " estimated random and total random errors?
Response: All the different metrics that are mentioned in this section were requested to be written here by the researchers involved in CrIS product/retrieval developments. Each of them is explained in Shephard et al. (2015), Shephard and Cady-Pereira (2015), Shephard et al. (2020) [all references can be found in the reference list of the manuscript]. All the metrics are explained there in detail and are also given in the product as shown in their metadata below.
If the reviewer and/or editor believe they are disturbing and/or not useful, we could remove some of the statistics:

```
netcdf Combined_NH3_p165_0_p180_0_n050_0_n045_0_20200101 {
dimensions:
        Observations = 1043 ;
        Layers = 15 ;
        RVMRLen = 5 ;
        nerr = 2 ;
variables:
        double xretv_meas_error(Observations, Layers, nerr) ;
                xretv_meas_error:units = "ppmv" ;
                xretv_meas_error:long_name = "Retrieved Species Measurement Error:
(Minus,Plus)" ;
…
        double rvmr_error(Observations, RVMRLen) ;
                rvmr_error:units = "ppmv" ;
```

rvmr_error:long_name = "RVMR Error " ;
…
double xretv_total_error(Observations, Layers, nerr) ;
xretv_total_error:units = "ppmv" ;
xretv_total_error:long_name = "Retrieved Species Total Error: (Minus,Plus)" ;
double tot_col_meas_error(Observations) ;
tot_col_meas_error:units = "molec/cm2" ;
tot_col_meas_error:long_name = "Retrieved Species Total Column Measurement Error" ;
double tot_col_total_error(Observations) ;
tot_col_total_error:units = "molec/cm2" ;
tot_col_total_error:long_name = "Retrieved Species Total Column Total Error" ;
…
double total_covariance_error(Observations, Layers, Layers) ;
total_covariance_error:units = "ln(vmr)^2" ;
total_covariance_error:long_name = "The total error covariance matrix is the sum of smoothing and measurement error (systmatic not included at this time). For atmospheric temperature, it represents the covariance of the error of temperature. For Atmospheric Species, it is the covariance of the error of ln(vmr)" ;
…
double noise_error_covariance(Observations, Layers, Layers) ;
noise_error_covariance:units = "ln(vmr)^2" ;
noise_error_covariance:long_name = "The Measurement error covariance matrix from the radiances. Presently it is also used as a lower limit on the Observation error covariance matrix (Measurement + systematic + cross-state errors) as systematic and cross-state are not explicitly derived at this time. The utility of the observation error is for comparisons with other measurements and for assimilation. The smoothing error is accounted for when one applies the averaging kernel, so the observation error accounts for everything else." ;

l 387: "2.410The calculated posterior uncertainty for our spatial domain and studied period (2013–2020) is shown in Figure 4Figure 4 (right)." Please correct.
Response: It has been corrected now (Track Changes P.14-L.431).

The new section 4.2 is quite long and may be condensed (to 1 page if possible) to become more in balance with the other sections. A few suggestions: The first paragraph is long, with some repetition, and may be reduced to a few lines. Fig. S8 could be removed and the second paragraph could be summarised in one line.
Response: We have tried to remove repetition from different parts of section 4.2 (see Track Changes P.18). Since this section justifies why this method is suitable for NH3 calculations and answers why a more classic method was not used, we would prefer to keep Supplementary Figures as in the previous version, because they are explanatory and – besides – they do not occupy additional space in the main manuscript.

For some stations the differences between the forward/backward calculations and LMDz are quite large, e.g. at DE0002R. Does this have implications for the emission

estimates? Could you provide a (rough) estimate of how much these modelling differences influence the uncertainty of the a-posteriori emissions?

Response: We have added a full explanation why this happens in Lagrangian simulations (see Track Changes P.20 – L.665-673). The main message, from the explanation given in the manuscript, is that backward simulations are always more accurate than forward ones, mainly due to the larger number of particles given per release (two more reasons are also described in P.20).

For example, in forward simulations, where there are hundred thousand of releases, the number of particles is weighted with the mass (the larger the mass, the larger the particle number), until an upper limit of 400 million particles is reached (larger number causes memory allocation problems). This set-up results in particle numbers of up to 1-2 thousand particles per release, at maximum. Backward simulations use 50 thousand particles per release (25-50 times more particles), due to the limited release points (one for each receptor), as usually releases occur at specific receptors (in the present case, at stations where we had observations).

Hence, no additional modelling uncertainty is expected in the posterior emissions of NH3.