

Interactive comment on “Estimating Upper Silesian coal mine methane emissions from airborne in situ observations and dispersion modeling” by Julian Kostinek et al.

Anonymous Referee #1

Received and published: 5 February 2021

The manuscript by Kostinek reports an interesting case study of top-down greenhouse gas emission quantification on the region scale using airborne observations, sophisticated atmospheric transport simulations and an inverse modelling framework. They estimate the CH₄ emissions from coal mining shafts in Upper Silesia, one of the CH₄ emission hot spots in Europe. The research topic is very relevant regarding greenhouse gas emission mitigation strategies under the Paris Agreement and as such deserves publication. The study design, the applied transport simulations and partly the inverse modelling approach are sound and mostly presented well in the manuscript. The authors took great care to assure adequate atmospheric transport simulations and made an important effort to characterise the connected uncertainties, adding some

C1

valuable new concepts to the field. There are two major concerns, 1) regarding the feasibility of assigning emissions to individual ventilation shafts given the flight observations taken at considerable distance from individual sources and 2) concerning the omission of CH₄ sources other than coal mines. These issues are detailed below and an answer will probably require additional analysis and revisions of the manuscript. However, I would encourage the authors to address these additional points, so that their very valuable analysis can be published in ACP.

Major comments

Inverse modelling method: I have two major concerns with the emission estimation method.

1) Emission attribution to individual shafts

The first concern is the attribution of emissions to individual facilities and shafts. Looking at figure 7 and 8 but also at the names of the shafts in figure 10, it is clear that many of the individual shafts cluster around individual mines at distances not much more than a kilometer. Furthermore, a lot of the locations actually line up with the main wind direction. In this situation it seems to be virtually impossible to estimate emissions from individual shafts from the presented observational data which was only taken at a single downwind curtain. The results presented in figure 10 are most likely a fine example of overfitting the observations and obtaining a ‘noisy’ a posteriori result. The problem is also apparent from figure 11, which seems to indicate that although there is some sensitivity to all emission shafts, sensitivity is much larger for certain shafts than for others. Given the observational data, the problem cannot be overcome in a general way, but at least the covariances in the inversion should be designed in a way that will limit overfitting and the overinterpretation of emission results. I would suggest that the authors modify the design of their a priori covariance. Currently, they only include diagonal elements. Hence, they assume uncorrelated uncertainties even for shafts from the same facility. I think it would be useful to explore by how much the results change if

C2

correlated a priori uncertainties for shafts from the same facility and/or shafts a shorter distances would be introduced (positive off-diagonal values in the a priori covariance matrix). Furthermore, the a posteriori covariance matrix should be explored in order to see if many of the emitters actually show negative covariances to one another. This would indicate that there remained a large uncertainty as to which shaft the emissions had to be assigned. In general, the manuscript should better highlight that the uncertainty on the shaft-level remained large and that the observational data is limited for a more specific estimate.

In this context, it would also be good to present the shaft level emissions on a map and compare with spatial inventories.

This issue was also raised in the comment by J. Necki and he provides valuable discussion on a per shaft basis that should not be ignored in a revision of the manuscript.

2) Neglect of other than coal mine emissions

In section 2 it is discussed that EDGAR assigns about 14 % of total CH₄ emissions in the area to non-coal mine emissions. The authors largely ignore these emissions in their analysis. Only in section 4.4 the possibility of emissions from the city of Krakow are discussed but ultimately these were also not included for the inverse modelling. Although 14 % may not seem a very large fraction and one could argue that any kind of mis-attribution is covered by the uncertainty estimates, it will still strongly depend on the distribution of these missing emissions relative to the sampling locations. Their impact on the concentration observations may well have been much larger if sources were closer to the flight track than the coal mine emissions. Since the authors also find that total emissions for the region are considerably smaller than what was reported in EDGAR, the fraction of non-coal may also be larger than in EDGAR. Seasonality of emissions may also play a role here. Since this is early summer, temperatures were above annual average, possibly leading to larger than annual average emissions from microbial sources like landfills, waste water treatment, manure management, etc.

C3

There seems to be a large discrepancy with EDGAR in terms of shafts and emission locations as well (red grid cell in the lower left corner of the encircled area close to Ostrava). Are these actually coal mine emissions in EDGAR? Looking at the area, one can also see a larger reservoir north-west of Bielsko Biala. Could natural emissions from this reservoir also play a role in the poorer agreement at the eastern end of the curtain flight? I strongly suggest that the authors reconsider their neglect of the non-coal mine emissions. They should obtain inventory data of these emissions, possibly not just from EDGAR (see comment below on inventories) but other more resolved inventories or from local information (see comment by J. Necki). With these emissions another FLEXPART forward run should be conducted and the resulting concentration either be removed from the observation vector before the inversion or an additional scaling factor for non-coal mine emissions should be introduced in the inversion.

Minor comments

p4: Regarding the use of the EDGAR inventory I would like to question if this is really the best available bottom-up inventory for the area. First of all, there is newer version of EDGAR available (v_5.0_GHG), which explicitly lists CH₄ from coal exploitation as a separate category and is available for a more recent year (2015) than EDGAR 4.3.2. Furthermore and as part of the EU project CHE, TNO has compiled higher resolution (6 km x 6 km) inventories for Europe. They may be better suited than EDGAR (see <https://www.che-project.eu/sites/default/files/2019-01/CHE-D2-3-V1-0.pdf>; data usually available on request). This is not only important for the final comparison of obtained emission estimates but also relates to the question if and how non-coal emission need to be treated in the inversion framework.

p5, l105ff: A detailed description of the flight data is given here. Some of it is depicted as well in figure 5, but it would be valuable to present all recorded time series of meteorological and CH₄ data along with flight parameters in a separate plot (supplement) highlighted by flight segment. One plot for each of the flights. If such a plot was already published previously a reference would be sufficient too.

C4

p5, l109: Here it is mentioned that an upwind concentration is subtracted from the downwind measurements. Later on a different method for background subtraction is described. What was really used?

p5, l113: Why is detrainment/entrainment important to this study? The FLEXPART-WRF simulations don't exclude detrainment/entrainment processes or PBL growth. Detrainment/entrainment would be more of an issue for a mass balance approach.

p5, l117: While an estimate of the morning PBL height is given, its height is not mentioned for the afternoon flight. Please add. Maybe also comment on the growth of the PBL height between the two flights and how this relates to the question of detrainment/entrainment.

p7, L148f: Were the Doppler soundings the only observations that were nudged? What about other standard synoptic observations in the area?

p8, l158ff: Does this mean that 3Dvar and nudging were applied to the same observational data? That would not make sense in my view as the same information gets used twice. Rather use 3Dvar with smaller error covariance if the pull of the observations seemed too weak and such smaller uncertainties could be justified. Also, how were observational error covariances determined exactly?

p8, l167: Are these numbers the root mean square errors between model and observations for the 1-Hz sampling?

p8, l171f: How can this apparent offset in pressure be explained? Difficult to believe that the models (WRF nested in GFS) are off by that much, especially since the wind seems to match very well. Was there any comparison to other surface pressure data? Concerning the temperature offset: Does this vanish when you calculate potential temperatures? And same question as for pressure: were there any ground based measurements to compare to?

p9, inversion method: I got confused by the description here. First, a non-regularized

C5

least square equation is presented for flux optimisation (eq. 1). Then regularization using a priori information and Bayes' theorem is advocated. To my understanding the resulting equations 4 and 5 only require a simple matrix inversion for solving for the a posteriori state. However, from line 211 onwards the application of a non-negative least square solver is presented. The latter is probably applied to equation 1, yielding a positive solution for x . However, if this was the case, I don't see why further analytical solutions to the cost function are presented in 4 and 5. I assume I am missing an important point here and would like the authors to clarify. If only a positively constrained solution for equation 1 is obtained I would think the results are even more overfitted as already mentioned above, since no additional a priori constraint on the individual sources would have been used. The description in the results section strongly suggests that this was the case. In equation, 4 I also think the last term Kx should also be Kx_a instead (see Tarantola eq. 3.37 or Jacob eq. 23).

p16, l336: Here the total uncertainty of the emission estimate is presented as the sum of the 'systematic' uncertainty (which I assume results from the a posteriori covariance; eq. 8) and the spread obtained from the sensitivity simulations. Why are these uncertainty terms not added quadratically?

Section 4.3: The way the covariance matrices for the a priori and the observation/model error are constructed most likely oversimplifies the true nature of the involved covariances and may lead to overfitted results. First, and already mentioned in the main comment above, the a priori covariance should acknowledge the fact that the a priori emission uncertainties will be correlated. This is true for shafts belonging to the same mining complex but may also be true for spatial distances between shafts. As mentioned above, I would suggest introducing off-diagonal elements in the covariance matrix to honor this fact. This would certainly lead to a smoothing out of the emissions across different shafts but is a more realistic approach. Furthermore, the observation/model covariance does not include off-diagonal elements either. However, the 1-Hz observations are certainly not independent from each other since they contain

C6

tempo-spatial autocorrelation. The latter will also be present in 1-Hz model residuals. I would suggest to explore this auto-correlation in the residuals and add a temporal correlation length to the observation/model matrix accordingly. Adding these off-diagonal elements will probably reduce the impact of the observations on the a posteriori results, reflecting that they are not really independent from each other. Another way to get rid of the autocorrelation would be temporal averaging of the observations before using them in the inversion. This has its merits as well as it would also bring the spatial resolution of the observations closer to those of the transport model.

p11, l231: How was the transport model uncertainty estimated concretely? As the standard deviation of simulated concentrations from the 8 ensemble members?

p12, l248: Why not use the measurements from the upstream flight segment as background? The comparison with the model output seems to indicate that the overall plume was wider than the flight segments.

p14, l296: This sentence largely repeats the result from the previous sentence (EDGAR being much larger than the current estimate).

section 4.6: After reading the first few sentences, it was not clear to me how an uncertainty quantified by sigma was adopted in the transport model. I guess figure 11 makes it clear that 8 sensitivity runs were done where the respective variables were perturbed in one or the other direction globally, but not perturbed by random noise with the given sigma width. This should be made a bit clearer from the beginning.

p15, l325: If I understand correctly, original horizontal wind speeds as output from WRF were increased/decreased by 0.9 m/s. In doing so, the local mass balance of the wind field may well be destroyed as vertical wind speeds were not adjusted (correct?). This may lead to errors in the transport description of the LPDM. Have you given this any thoughts? Probably the impact was not to large and since this is only presented as a sensitivity case it is of less importance, but it may have lead to larger discrepancies from the reference run than anticipated. A similar question for the PBL height. Is the

C7

latter taken from WRF or is the diagnostic calculation taken from FLEXPART? When increasing the PBL height just in FLEXPART vertical mixing in FLEXPART may then bring model particles to altitudes that in WRF are not part of the PBL and as such may have a distinctly different flow direction as flow in the PBL. As a consequence the differences to reference run may be larger than in a case where WRF PBL heights were larger/smaller. Hence, your change in the PBL height may give a slightly more pessimistic (larger) uncertainty.

p16, l23: Here it is mentioned that the statistical uncertainty was estimated from eq. 7 and 8 and it is referred to elements e_j , which are the diagonal elements of the a posteriori covariance matrix. What about the off-diagonal elements of this matrix? Were the taken into account for the total uncertainty estimate?

Figure 12 and use of Jacobian: If I understand correctly, what is shown in figure 12 is the matrix K containing the elements dy_j/dx_i . However, the term Jacobian is also used in the manuscript for $\text{grad}(J(x))$. But J is not a vector-valued function and as such $\text{grad}(J)$ is not a Jacobian. Please clarify.

Technical comments

p2, l45. Karion et al. is missing a publication date.

Figure 1: The line indicating the afternoon flight is more orange than red (as described in caption). It looks like the map is showing total EDGAR emissions. How does the distribution of non-coal mine emissions look like?

Figure 3: Please label the WRF domain in the figure according to their definition in the text.

Figure 4+5: Please use the same colors for the different WRF runs. The legend is fairly small in both cases and needs to be enlarged, possibly put to the right of the sub-panels.

C8

