

Comment on egusphere-2026-3665

This paper describes an inverse modelling approach developed for use with aircraft data, with explicit optimisation of the background within the inversion framework. The approach is demonstrated by applying it to a single flight around Berlin. It is found that this inverse modelling based approach yields more a more reasonable emission estimate, with a lower uncertainty, than a mass balance approach using the same data. I agree that the mass balance method is too crude to apply to most urban areas and I also agree that the mole fractions of the inflowing air need to be optimised within the inversion. However, I have some questions regarding the approach taken here that I think need some extra thought.

My main issue is that I don't really understand the motivation behind this particular construction of the state vector. As far as I understand it, in this case it contains 5 emission scaling factors and 15 boundary condition scaling factors. The emission scaling factors are assumed to be uncorrelated (if I've understood that correctly then it should be stated explicitly in section 2.5.1), while the background offset correlation is adjusted so that the background concentrations don't become more correlated in the posterior. The rationale for this last part seems unclear to me. It is true that it would be undesirable to push the plume into the background offsets, and that this would lead to a higher correlation in the background. But an upwind plume that should be in the background (but is not well captured by the prior) would surely have the same effect. So increasing the background correlation in the posterior does not seem inherently bad.

Previous studies (e.g. Lopez-Coto et al., 2020) have handled the background by optimising the flux field at the grid scales (with appropriate spatial correlations) in a wider domain. This makes good use of the upwind transect – the inversion can use this data to optimise the upwind flux map, and this optimised background is also propagated onto the downwind transects. Of course there is still some risk that the inversion responds to a large downwind city plume by increasing sources far upwind, but the upwind transect acts against this (especially if the model-data mismatch errors have an element that is proportional to the magnitude of pollution and/or the variability within the averaging period). It seems like this spatial information regarding the background is essentially lost in the method presented here, and that is a shame because it is one of the big advantages of flying the aircraft. I'm not saying that the authors should necessarily adopt the approach of Lopez-Coto et al., 2020, but I think this element of the inversion needs more thought, especially as the treatment of the background within the inversion is the main focus of this paper.

One potential issue with employing a more spatial approach to the background optimisation is the limited amount of data. This raises two other questions for me. Firstly, why were footprints calculated every 5 minutes? At the speed the aircraft flies it seems to me like you would get considerably more information by releasing footprints every 1 minute. I know that adds to the computational expense, but that is nothing compared to the expense of flying the aircraft, and if you only take a point every 5 minutes I feel again that these aircraft data are not being used to their full potential. The second question I have concerns the decision to "exclude measurements that exhibit negative prior correlations in background concentrations". This needs much clearer explanation and justification – at the moment I don't understand how this filter works. It would also be good to state

explicitly how many measurements are there before filtering. When I look at Figure 1 it seems like there are hardly any remaining red crosses on the upwind transect, which poses an issue.

On a separate note, it would be useful to have more information on how the hour-in-day scaling factor was applied to the TNO inventory in the prior, and how the time of day was dealt with in the VPRM run. The flight itself spans the middle of the day, but what was the timeframe for the sampled air passing over Berlin? This seems especially relevant when making comparisons between the posterior and the inventory in the discussion section – the net biospheric flux changes rapidly during the morning from net emission to net uptake. As a related point, more information is needed on the VPRM configuration, especially regarding changes made for urban areas. This should include reference to Li et al. (2026), even if a different framework has been used here.

L6 – “using *the* Weather Research and Forecasting (WRF) model”

L103 – “the concentration data *for* the background field”

L111 – check citation format

L144 – I would say that background values are just as often interpolated between the edges of the downwind transects. In any case both approaches are in common use so it probably makes sense to mention them both here.

L165 – More clarity with the terminology would be good here. If x contains emission scaling factors then presumably the various components of f are not footprints (at least according to the standard definition) but a multiplication of the footprint by the emissions in each grid cell?

L187 – “gain and *averaging* kernel”. Also I’m not sure that these are usually capitalised?

L190-192 – should x_0 be x_a here for consistency with above?

L265 – the reference to “the sensitivity study” here is ambiguous – be explicit that this is an attempt to quantify uncertainty in the posterior related to uncertainty in the specification of the prior covariance matrix.

L364 – repeated sentence

References:

Israel Lopez-Coto, Xinrong Ren, Olivia E. Salmon, Anna Karion, Paul B. Shepson, Russell R. Dickerson, Ariel Stein, Kuldeep Prasad, James R. Whetstone; Wintertime CO₂, CH₄, and CO Emissions Estimation for the Washington, DC–Baltimore Metropolitan Area Using an Inverse Modeling Technique. *Environ. Sci. Technol.* 3 March 2020; 54 (5): 2606–2614. <https://doi.org/10.1021/acs.est.9b06619>

Li, J., Chen, J., Glauch, T., Brunner, D., Marshall, J., Ponomarev, N., et al. (2026). Fine-scale estimation of urban biogenic CO₂ fluxes: A novel framework integrating multiple versions of vegetation photosynthesis and respiration models and in situ measurements. *Earth's Future*, 14, e2025EF007458. <https://doi.org/10.1029/2025EF007458>