Authors' response to review of egusphere-2025-1568

We thank both reviewers for their overall positive and constructive feedback on the manuscript. In addition, we would like to thank Delong Zhao for additional feedback through a community comment. We have considered all comments and have revised and improved the manuscript accordingly.

Shortly after submitting the preprint, we discovered a bug in how the background concentrations were computed (it only accounted for 14 days of transport rather than 30 days). With this in mind, we decided to redo the model simulations and at the same time take into consideration some of the major concerns raised by the reviewers, e.g. accounting for negative emissions and extending the sensitivity analysis. The overall spatial and temporal pattern of the posterior is unchanged, but the magnitude is closer to the prior: +13% compared to +18%. The annual posterior uncertainty is also increased from previous estimated 10 Gg to 16.4 Gg, mostly due to the inclusion of site-specific MAC sensitivity.

Following the main points raised by the reviewers and the changes to the posterior and the sensitivity analysis in the new version, we have modified the overall discussion and conclusions to better reflect the underlying assumptions and uncertainties.

We have also taken the opportunity to make some minor grammatical changes in the revised manuscript, but they do not change the meaning of the content.

Reviewer #3

'Comment on egusphere-2025-1568', Anonymous Referee #3, 07 Nov 2025

General comments:

Summary of paper

This paper presents a top-down approach to estimate anthropogenic black carbon (BC) emissions in Europe in 2021 using a Bayesian inversion framework (LUMIA) coupled with the FLEXPART dispersion model. The paper is well organized and provides a thorough literature review with clear comparisons to previous studies. The authors find that the top-down (posterior) estimates are approximately 18% higher than the bottom-up (prior) inventory. Through detailed spatial and seasonal analyses, they identify regions where bottom-up inventories likely underestimate emissions. The study also includes comprehensive sensitivity and uncertainty analyses, offering valuable insights for improving emission inventories and informing future policy and modeling efforts.

With some minor edits, this manuscript should be published.

We want to thank reviewer #3 for the positive and thorough review which has helped us to improve the manuscript. Below you can find our point-by-point response to all comments.

Specific comments:

Line 76: You mention the error term but you don't explain how it was calculated. Adding a brief sentence or two describing how it was calculated would help.

We have updated the description of the error terms, including separating the explanation for the observational error and model error:

"The observational error, ε_o, is further described in Sect 3.2. The true model error, ε_m, is unknown, instead we estimate it as site-specific weekly uncertainty, based on the difference between the modelled and observed short term variability (see Monteil et al, 2025, for further details)" (Ln 76)

Line 101: "Calculating footprints backwards in time is numerically more efficient in our case, since we have fewer observational sites than grid cells"

Can you elaborate more on this? I'm not exactly understanding why this would be more numerically efficient.

With FLEXPART we essentially relate how emissions in every source grid cell affect the concentration at each site (often called Source Receptor Relationship (SRR) or footprints (FP)). If we would compute this forward in time, we would need one FP release (simulation) per spatiotemporal-gridcell while in backward mode we only need one release per time at each site. The underlying intuition for this is that, in forward mode, we would compute much more information than we need, essentially the SRR for each cell to every other cell.

Line 213:

"First, by changing the horizontal correlation length Lx from 500 to 250 and 1000 km (SCx.), and then by changing Lt from 14 to 7 and 21 days 215 (SCt.)"

Can you clarify this sentence? I'm a bit confused by the three numbers for Lx and Lt. What are you changing the numbers to and from?

We have clarified to: "First, by changing the horizontal correlation length, L_x, from 500 km in the synthetic base case to 250 km (*SCx.205*) and to 1000 km (*SCx.1000*). Secondly, by changing the temporal correlation length, L_t, from 14 to 7 days (*SCt.7*) and to 21 days (*SCt.21*)" (Ln 213)

Grammar:

Line 69: should be "detail"

Fixed

Line 80: there should be "the" before cost function

Fixed

Line 95: there either needs to a new sentence starting at "Useful" or you need something before it like "thus it is".

Fixed

Line 155: you have "the" twice

Fixed

Line 188: remove comma after "quality controlled"

Fixed

Line 288: "to winter" should be in "in winter"

Fixed

Line 333: “averaged” to “average”

Fixed

Line 449: Black Carbon needs to be lower case

Fixed

Line 449: change “This have” to “This has”

Fixed

Line 450: Please clarify or reword this sentence: “less emissions in climate or atmospheric transport models may lead to underestimating of for example radiative forcing or air quality effects.”

Changed to: “using underestimated emission inventories in climate or atmospheric transport models may lead to underestimated radiative forcing or air quality effects. Considering both top-down and bottom-up emissions in modelling efforts may provide more robust outcomes and help quantify biases and uncertainties.” (Ln 450)

Line 451: This sentence also needs to be redone: “In addition, this work is a first step in identifying what (missing or misrepresented) sectors in bottom-up inventories is the driver for this underestimation, information which can help guide future policy changes”

Change “identifying what” to “identifying which”. The last part after the comma can also be improved if you add “which provides information that can help..” or something like that.

Changed to: “In addition, this work is a first step in identifying which (missing or misrepresented) sectors in bottom-up inventories are the driver for the apparent underestimation, which provides information that can help guide future policy changes” (Ln 451)

Review #4

'Comment on egusphere-2025-1568', Anonymous Referee #4, 23 Nov 2025

Overall Assessment

This study presents a valuable effort to refine estimates of black carbon emissions over Europe by applying a Bayesian inversion framework to surface observations. The topic is of clear importance for emission mitigation policies. However, the reliability of the conclusions requires stronger substantiation, and major revisions are necessary to address several key concerns.

We want to thank reviewer #4 for the thorough review. It has helped us to improve the methodology and manuscript. Below we have responded to all comments point-by-point.

Major Comments

1. A primary concern is the attribution of nearly all discrepancies between observed and simulated concentrations to emission errors, without sufficiently accounting for other potential sources of systematic bias. These include errors in the transport model, uncertainties in the conversion from light absorption to equivalent black carbon (e.g., the mass absorption cross-section), inaccuracies in background concentration estimates, and biases in wet/dry deposition parameterizations. The inversion framework appears to assume these errors are either negligible or adequately represented in the observation error covariance matrix R, yet no validation is provided for this critical assumption. Consequently, the extent and spatiotemporal structure of the reported "emission underestimation" may be significantly overstated.

Accounting for systematic biases is a fundamental issue in any inversion system, since the underlying uncertainties in the Maximum A Posterior optimization is assumed to be random errors. A weak-constrain 4DVAR (that could account for systematic biases by applying an error term to the cost function) is currently not implemented in LUMIA, therefore, there is no way for the optimization to handle systematic biases directly. These biases can instead be investigated with sensitivity analysis as we have done for numerous potential biases (both in the original preprint and in this response; testing for different MAC values, see later review comment). Considering the results from the new sensitivity tests and the discovery of a bug in the calculation of the background concentrations, we agree with the reviewer that the discussion and conclusions of the preprint does not reflect the underlying systematic biases and uncertainties well enough. This is changed in the new version, e.g:

"The results show that the bottom-up BC inventory on average underestimates emissions in the domain, given our assumptions regarding transport, background concentrations and mass absorption cross-section (MAC)." (Ln 7)

"However, the inversion shows high sensitivity to the assumed MAC of 10 ± 1.33 m2g-1 at 637 nm. We also assume the transport model to be representative of the actual transport and in extension the computed background concentrations to be correct.

Under these assumptions, the posterior annual emissions from the inversion with the real observations is found..." (Added Ln 430)

We do want to point out that the observation error covariance matrix does consider a site-specific transport model error based on the difference short-term obs-model mismatch and instrument and MAC uncertainty, often referred to as "error inflation". It aims at reducing the cost of departing from a single observation close to what it would have been if obs error correlations had been accounted for. The description of the observational error has been updated in Sect. 2 to clarify this.

2. The processing of observational data and site-specific quality control procedures lack sufficient detail, raising concerns about potential systematic biases. The manuscript would benefit from a clearer description of the quality control protocols, outlier handling, the sources and uncertainty propagation of instrument-specific correction factors (such as for the AE33), and the steps taken to ensure comparability across different instruments and sites. Given that measurement errors for black carbon are often systematic, failing to properly quantify and propagate these uncertainties risks misinterpreting observational biases as emission signals.

We exclusively use Level-2 quality controlled light absorption coefficients observations from EBAS, which undergo well documented quality control protocols and outlier handling (Müller and Fiebig, 2021). We assume that the reviewer does not expect us to include documentation of this QC in the manuscript. Generally, the point of ACTRIS and other observational network infrastructures hosted on EBAS is that the Level-2 quality controlled data should be usable without additional QC. That said, we have updated Sect. 3.2 to better reflect some QC choices we did which include: manual inspection of the timeseries, site selection, instrument selection for sites with multiple instruments and application of correction factors:

"We performed an additional quality control of the Level-2 data. First, the timeseries were manually inspected for outliers (none found). Then, data selection was performed by 1) only selecting rural or semi-rural (two sites) sites to remove influence of local (e.g. street-canyon) emissions not resolved by the model and 2) in cases of multiple collocated instruments selecting firstly for the MAAP if available or else the one with the most data coverage. Finally, we manually applied a correction to specific sites (see description of the Aethalometer AE33 below)" (Added Ln 164)

One of the major concerns that we discovered was the use of the filter specific correction factor, H*, supposed to be applied to measurements from the AE33. The purpose of this is to align the AE33 with MAAP measurements and it is the recommended practice from ACTRIS guidelines (Müller and Fiebig, 2021). The aethalometer consistently underestimates mass absorption coefficients (overestimate eBC). The main issue for us was that the usage of this correction factor was inconsistent for 2021 and inconsistently mentioned in the metadata. We discovered this when comparing mass absorption coefficients from collocated AE33 and MAAP measurements at *hbp*, where it was clear that the factor had not been applied and not mentioned in the metadata. This triggered an investigation into the other sites with only AE33. At three sites (*jfj*, *dem*, *htm*), H* had been applied and documented in the metadata. By contacting PIs, we found that for three stations (*kos*, *rig*, *pay*) H* was not applied and for one station it was applied but not documented (*vda*). We were unsuccessful in contacting the PI at the one site (*pmi*). Here we utilized the prior forward run, which underestimated observations by about a factor of 2, to conclude that the factor had not been applied. The same factor 2 underestimation in the prior forward run was also observed in the other three cross-references site that did not apply H*.

We have updated the manuscript with a more detailed description of this process, which ultimately aimed to increase the cross-instrument comparability of the assimilated observations, one of the reviewer's major concerns:

“By comparing collocated AE33 and MAAP at *hpb*, we found that the application of this factor, along with accompanying metadata, was inconsistent for 2021. Only three out of eight AE33 sites (*jfj*, *dem*, *htm*) included information about that the usage of this factor in the metadata (it had been applied). To remove this bias in level-2 quality controlled AE33 measurements, we cross-referenced with data providers and found that the factor (1.76) had not been applied at *kos*, *rig* and *pay* while it had been applied at *vda*. The cross-reference was unsuccessful at *pmi*, but forward model simulations showed a factor of around 2 underestimations at that site from which we concluded that it had not been applied there. Finally, we manually applied the harmonization factor to the data from the four sites that had not implemented it.” (Extended Ln 184-187)

3. The validation of the inversion results relies heavily on the improved fit to the assimilated observations and the comparison with the prior emissions. There is a lack of validation using truly independent evidence, such as external observational datasets, independent emission inventories, or evaluations with independent model simulations. Without this, features like the reported seasonal emission enhancement in Eastern Europe could arise from model artefacts or sampling biases rather than representing true emission patterns.

While we understand the reviewers point, we do not fully agree with the criticism. It is true that the manuscript includes comparison of the observational fit between the prior and the posterior for the dependent observations. This is purely for diagnostic purposes – showing that the inversion converges to a solution with improved fit.

We do argue that the cross-validation scheme discussed in section 5.2.3, which is validation of the robustness of the inversion with independent observation, is a stronger validation than the alternative of leaving out a few stations for independent validation. The latter approach suffers from the same issues the reviewer points out. There is a potential risk that whatever stations/observations that are not included in the inversion introduces sampling biases in the posterior. This is certainly the case when the network is spatially sparse and temporally confined to a single year. Figure B2 shows exactly this; if we would have left out for instance Pallas for validation a large positive bias in the posterior would have been introduced. In addition, validation against independent observations can suffers from the same bias as the assimilation of similar data, e.g. if the transport model overestimates the sensitivity to emissions for the data it assimilates, it will probably do the same for the validation data,

With the cross-validation approach, we include all stations, some of which provide crucial spatial information, and at the same time remain confident in the robustness of the posterior. The cross-validation shows that the inversion does converge to similar solutions even if the network changes slightly and that this solution results in an improved match against independent observations in most cases. In addition, we can quantify the uncertainty introduced by a single site and analyse the response of the inversion when nearby stations have conflicting signals. We have modified the text in Sect 5.2.3 to more clearly communicate our point of view:

“To validate the posterior, independent observations which are not used in the optimization should be used. However, each observation left for evaluation is an observation not assimilated in the posterior, which can introduce sampling biases. Therefore, we opted for a cross-validation scheme with multiple inversions, leaving one site for validation each time. With this approach we evaluate the robustness of the inversion to changes in the network and evaluate how an entire ensemble of similar inversions performs on independent observations. We do acknowledge that this approach does not validate the exact posterior where all sites are assimilated, i.e. the real base inversion (*RBase*), but we argue that the cross-validation provides enough confidence in the inversions robustness while still utilizing all available observations. It also indicates how the base inversion is expected to perform on new observational data, given their area of influence." (Ln 326-330)

To further illustrate the robustness of the inversion system, we here present a new inversion where we have utilized gaps in the data to generate an independent dataset for testing. For each site we locate two 7-day periods of observation in contact with the first and second largest gaps in the data. The rest is kept for assimilation, except 3.5 days of observations on either side of these two test periods to remove potential temporal autocorrelation. Figure 1 illustrates this split for all sites. In total, 36539 observations are assimilated, 1786 are for testing and 1126 are discarded. The posterior performs on average across the stations better on the test dataset than the prior. We find a 10% improvement in nRMSE (from 0.27 to 0.24), a 15% improvement in correlation coefficient (from 0.51 to 0.58) and a 35% improvement in nME (0.031 to 0.020).

Crucially, the posterior generated in this experiment is very similar to the posterior of the RBase inversion both in magnitude (-0.7% smaller) and in spatiotemporal pattern (not shown). We argue that this, together with the cross-validation scheme, shows that the RBase inversion indeed is robust to independent observations. Therefore, we do not feel it necessary to include this analysis in the final publication.

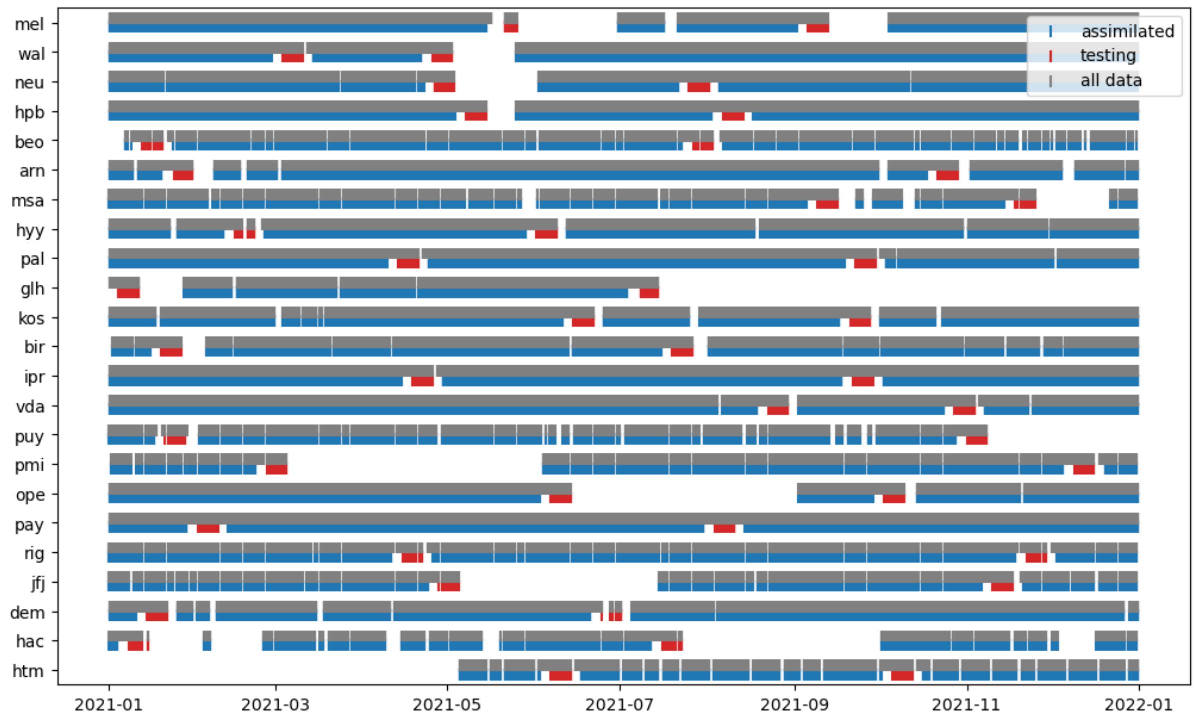

Figure 1: data coverage for all stations (gray), split into assimilated (blue) and test sets (red)

In future work, including a 10-year BC inversion, we plan to include both validation by left out observations and independent model evaluation. But this is outside the scope of the current manuscript. To our knowledge, there are no published BC top-down emission inventories of the same domain and year that could be used as independent validation. We do use the, to our knowledge, closest published dataset for comparison (Evangeliou et al., 2021) as validation in the discussion.

4. The conclusions exhibit a strong dependence on several choices—including prior covariance length scales, key optical parameters (MAC, AAE), and data selection criteria (e.g., time-of-day sampling). However, a systematic sensitivity analysis of these factors is absent. Providing such an analysis is crucial for readers to assess the robustness of the findings.

We agree that the conclusions are dependent on several choices, but we do not agree that systematic sensitivity analysis of these factors is absent.

The manuscript does include sensitivity analyses for several inversion settings in both our real and synthetic experiments, including prior covariance length scales. See Figure 2 and Appendix B1. We concluded that the inversion is not sensitive to these settings, only varying -3.5% to 2.5% from RBase.

For data selection criterion, our cross-validation experiment provides a sensitivity towards selection of observational sites, see Appendix B2. Sensitivity towards time-of-day sampling is not included in our manuscript for three reasons. 1) this method is the standard method utilized in many FLEXPART inversions papers because 2) previous work has shown that FLEXPART performs worse when the boundary layer is developing so with observations outside the selection, we risk introducing avoidable model errors. 3) At the same time, other preliminary work (Annadate et al., 2025) indicates that BC inversions with FLEXPART in Europe are not sensitive to this selection criteria.

We did not include any sensitivity tests toward key optical parameters. Evangeliou et al (2021) tested the sensitivity of a BC inversions in Europe towards MAC values by using a constant 5, 10 and 20 m^2/g and site-specific MAC values. They conclude that the overall sensitivity of the posterior is less for MAC compared to using different prior.

In addition, a recent literature review found an average MAC at 550 nm from filter photometers of 11.6 m^2/g across many articles (Asmi et al., 2025), which adjusted to 637 nm is 10.03 m^2/g, very close to our assumed MAC value. This indicates that our choice of MAC value is indeed the most appropriate choice. However, Asmi et al (2025) report overall higher uncertainty in MAC (3.2 m^2/g at 637 nm for all ambient measurements) than the one we assume (1.33 m^2/g from Zanatta et al., 2016).

Still, it is clear that there is a sensitivity towards MAC (Evangeliou et al., 2021). Therefore, we have extended the analysis by performing inversions with constant MAC values (at 637 nm) of $\pm 2\sigma$ of our assumed MAC of 10 ± 1.33 m^2/g. i.e. 12.66 and 7.34 m^2/g. This effectively means that we test the sensitivity towards a systematic bias with decreased and increased observed concentrations. Both values are within the reported values from Asmi et al. (2025). In addition, we test 50 sets of inversions where each stations MAC value is randomly sampled from a normal distribution of our assumed MAC (with max/min of $\pm 2\sigma$), to simulate station specific MAC values.

The results show a strong sensitivity towards MAC values. With a MAC value of 7.34 ± 1.33 m^2/g at 637 nm at all stations the total posterior emission in the domain is 528 Gg (34% higher than the posterior for base inversion), while a MAC value of 12.66 ± 1.33 m^2/g results in a posterior of 313 Gg (20% less than the base posterior). This illustrates, as the reviewer highlights, that the final posterior is highly sensitivity towards biases in the assumed MAC values. This will be the case for any BC inversion using filter based optical measurements of eBC. However, as stated previously, multiple sources highlights (Asmi et al., 2025; Zanatta et al., 2016) that a MAC value of 10 m^2/g at 637 nm is indeed the most appropriate choice for European sites, and we find it unlikely that the true MAC value for all assimilated sites is either as high as 12.66 or as low as 7.33 m^2/g.

"However, deceasing (increasing) MAC uniformly results in a large deviation of +34% (-20%) highlighting the sensitivity towards choice of MAC value. At the same time, we argue that [these experiments] is unlikely to represent the truth. For instance, the prior fit to observations is worse than *RBase*. It is more likely that there are biases in site specific MAC values (discussed later). In addition, multiple sources highlights (Asmi et al., 2025; Zanatta et al., 2016) that a MAC value of 10 m^2/g at 637 nm is indeed the most appropriate choice for European sites" (Added Ln 314, then moved to 296)

The 50-sensitivity test with random site-specific MAC allows us to quantify the uncertainty introduced by site specific MAC biases. Overall, the total posterior emissions in these simulations range from 373.3 to 433.7 Gg, with a standard deviation of 12.8 Gg. Notably, this uncertainty is higher than the posterior uncertainty presented in the preprint, which combined both uncertainty from perturbed prior emissions, observations and cross-validation scheme. Combining all three results in a final uncertainty of the posterior domain total yearly emission yields 16.4 Gg.

"Finally, the posterior uncertainty to MAC is quantified by perturbing site-specific MAC values, which introduce most variability near sites. This results in the largest annual standard deviation of 12.8 Gg, again highlighting that the inversion is more sensitive to the assumed MAC than any other tested parameter. Despite the inclusion of several uncertainty factors, the combined annual uncertainty of the posterior is estimated to be 16.4 Gg a reduction of 45% from the prior. Spatially, the combined uncertainty results in an overall reduction of uncertainty in the central part of the domain, where the observational network density is high." (Extended Ln 439)

In all MAC sensitivity tests, we assume that the absolute observational uncertainty remains the same as in the real base inversion. Otherwise, we would not only test the sensitivity towards MAC values, but also the observational uncertainty. Sensitivity towards eBC observational uncertainty is already included in the manuscript. However, we have updated the discussion to better relate the results from that sensitivity test to assumptions made regarding MAC standard deviation:

"The sensitivity towards the standard deviation in MAC is not tested explicitly. However, the sensitivity towards the observational uncertainty (which is derived partly by the MAC standard deviation) shows less sensitivity than other settings. This indicates that our assumption of MAC standard deviation is less important than, for instance, correlation lengths or biases in MAC values." (Added before Ln 315)

The focus of the manuscript is for BC, for which an AAE = 1 is generally a good approximation assuming bare soot (Moosmüller and Arnott, 2009). However, it is true that AAE of BC can vary depending on size and shape of the particle (Helin et al., 2021; Romshoo et al., 2021). In addition, Asmi et al., (2025) tested this assumption and found that MAC at 550 nm from all included studies only vary between -4% and +7% when varying AAE from 0.8 to 1.3. Since these changes are less than the changes to MAC in the sensitivity tests described above, we argue that testing the sensitivity towards AAE is not necessary in the current manuscript.

5. The interpretation of identified features, such as emission hotspots and seasonal patterns, remains somewhat qualitative. A more in-depth discussion of the potential physical mechanisms—for instance, linking patterns to boundary layer dynamics, specific source

activities (like agricultural burning or residential heating), or regional transport pathways—would greatly strengthen the discussion and provide deeper insight.

As clearly stated in the introduction, the main purpose is to test and validate the system. Therefore, we decided to only include a qualitative discussion of the physical mechanisms of the prior-posterior difference. As we also state in the discussion, we are not comfortable in drawing qualitative conclusion since only a single year is analysed.

However, we do agree with the reviewer that the analysis of potential physical mechanisms would provide deep insights. Therefore, a 10-year BC inversion with the current system is currently on-going and the resulting posterior will be used to draw more qualitative conclusions on the driving mechanisms. This is also stated in the discussion, but we have changed the wording to provide better context:

"However, this [possible causes for the difference found] is speculations based on a single year. Further investigation on a longer time-period is required to draw quantitative conclusions about the physical mechanisms behind the mismatch between prior and posterior. Therefore, future work will include a multi-year inversion with the inversion setup described and validated in this study." (Added Ln 368)

6. Notably, the real base inversion reported in the manuscript includes "negative emissions" accounting for 0.6% of the total annual emissions, a result that is physically implausible. The authors have not implemented any constraints to address this issue, which undermines the physical consistency of the inversion outcomes.

We are aware of this flaw in the system, and since the posting of this preprint we have implemented a solution for the negative emissions. As in Bergamaschi et al., (2022), a 'semi-exponential' description of the PDF for emissions is applied to enforce positive posterior emission. For a given control vector element, x, the emissions, E, in corresponding cell is computed as

$$E = \begin{cases} Eb * (1+x) \; for \; x >= 0 \\ Eb * exp(x) \; for \; x < 0 \end{cases}$$

where Eb is the prior emission in the same grid cell. The control vector is still 0 a priori, and we still assume it to be Gaussian (Bergamaschi et al., 2022). This assumption will introduce a slight bias, as the prior uncertainty for any negative state vector values should be lognormal. However, as shown below, the inversion with this semi-exponential mapping converges to a solution which is very similar to the original setup and the fit to the dependent observations are as good. This shows that the bias is small. The original LUMIA setup optimizes for an offset E = Eb + x, which is the absolute equivalent to only the upper part of the equation, E = Eb(1+x) (optimizing for a scaling factor).

Here, we compare the semi-exponential description with both the offset and scaling factor optimizations. Both linear methods result in nearly identical fit to the dependent observations with a network average nRMSE = 0.075, r = 0.744 and nME = 0.011. The semi-exponential (SE) fit is very similar with nRMSE = 0.078, r = 0.730 and nME = 0.009. Comparing the posterior emissions, we find that the linear methods result in a domain total posterior of 393 Gg/year with 2.5 Gg of negative emissions (0.6% of the total annual emissions). The SE approach results in total posterior of 394 Gg/year with zero negative emissions. Figure 2 shows the domain total daily emissions for linear methods in blue and orange, and SE in green. We see that the two linear methods are identical, while the SE only differs slightly. Spatially, we find that the SE compensates for less emission reductions by applying less emission increases (not shown for brevity). With nearly equivalent fit to the observations and posterior solutions, we argue that this is a viable solution to the problem of negative emissions *in this case* and have implemented it in the final manuscript.

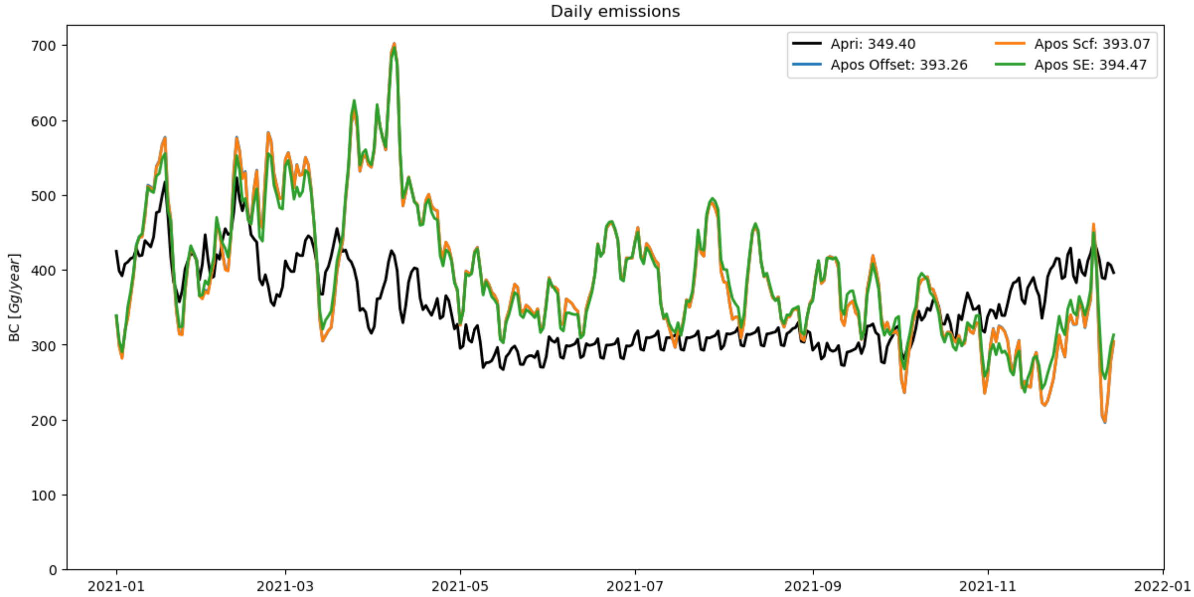

Figure 2: Daily emissions of the prior (black), the two linear approaches in LUMIA offset (blue) and scaling factor (orange) and the newly implemented semi-exponential description (green)

We have updated the description of the inversion framework to include a section that describes the SE approach (which is used in the new base simulations):

"Previously, inversions with LUMIA allow for negative values in the solution. This is strictly non-physical for BC assuming correctly modelled deposition. To solve this, we implement the 'semi-exponential' (SE) description of the probability density function of the emissions from Bergamaschi et al. 2022). For a given control vector element x the corresponding emissions, E, are computed as [the equation above] where Eb is the corresponding prior emissions. This means that we are optimizing for a SE scaling factor. Apriori, the control vector is zero and assumed to be Gaussian (Bergamaschi et al., 2022). Despite this assumption, as we will show, this approach converges to a solution very similar to the two previously implemented approaches in LUMIA (both linear). These are either to optimize for an offset (Monteil and Scholze, 2021), where E = Eb + x or a scaling factor (Monteil et al., 2025) computed as E = Eb (1 +x). Note that the scaling factor approach is the linear version of Eq. [equation above]. In both the scaling factor and SE approaches the control vector is dimensionless. (Added Ln 91)

And we have included a comparison to both the original (offset) and relative linear mappings (scaling factor) in the sensitivity tests:

"Finally, we test the sensitivity towards the choice of mapping from control vector to emissions (i.e. what we are optimizing, see Sect 2.1) by changing from semi-exponential scaling factor to linear scaling factor (*SM.scf*) and to linear offset (*SM.ofs*)." (Added Ln 216)

Community comment #1

'Comment on egusphere-2025-1568', Delong Zhao, 22 Sep 2025

General comment:

This study addresses significant discrepancies in European black carbon (BC) emission inventories by employing the LUMIA inversion algorithm and FLEXPART dispersion model to assimilate observational data from 24 background sites during 2021, thereby providing the top-down estimates of anthropogenic BC emissions at an annual scale. The study's innovative high-resolution inversion of European BC emissions is commendable, but flaws in background concentration treatment undermine its core conclusions. The claim of "18% underestimated European BC emissions" lacks credibility unless critical methodological gaps are resolved.

We want to thank Delong Zhao for taking the time to read and review our manuscript and thereby helping us to improve it. Below you can find our response to your comments.

Specific comments:

Line 109: The settings for wet deposition efficiency (rainout/cloud scavenging) significantly impact BC lifetime (4-10 days), yet no tests evaluate how different parameters (e.g., snow scavenging efficiency) affect posterior emissions.

The FLEXible PARTicle (FLEXPART) dispersion model has too many assumptions. How did the authors address this issue?

We agree that FLEXPART, as any other transport model, depends on many assumptions. However, FLEXPART is one of the most well documented, utilized and validated Lagrangian transport models that has been used in numerous studies for numerous purposes including BC inversions studies (Evangeliou et al., 2021, 2018; Jia et al., 2021). Specifically for wet deposition efficiency, not only are the values recommended and validated in FLEXPART development papers (e.g Grythe et al., 2017) but they have already been tested for BC inversion purposes (Evangeliou et al., 2018).

Lines 128-130: The manuscript claims that "total emissions are irrelevant for background calculations; only the ratio between intra- and extra-domain emissions matters." While theoretically valid under proportional scaling assumptions (e.g., doubling both intra- and extra-domain emissions preserves their relative contribution ratio to background concentration ybg), this contains serious flaws in practical application. The authors tested the impacts of prior uncertainties and observational errors but did not test the sensitivity of ybg to extra-domain emission errors. The manuscript simultaneously states "all emission sources are assumed to have identical mass absorption cross-sections (MAC)" (Line 169), yet MAC values for extra-domain BC (e.g., dust-influenced regions) versus intra-domain BC (e.g., European vehicular emissions) differ significantly, further invalidating the proportionality assumption. Line 405 indicates: "At most, 52% of the average concentration is attributed to ybg at Jungfraujoch (jfj) and Pic du Midi (pmi)." The poor performance of these sites in cross-validation (Figure 7) directly reflects the unreliability of background concentration calculations. If ybg contains biases, the system will misattribute external errors to European emissions, rendering the results invalid. The reported "18% underestimation of European BC emissions" may partly compensate for extra-domain emission/transport errors rather than reflect actual missing sources. How did the authors address this issue?

We agree with the comment, but we also need to limit the scope of the article by making some assumptions. One of these are the constant AAE = 1 (e.g. same MAC for all sources). See our answer (final paragraph) to Reviewer #4, point 4. However, the discovery of the bug in the background calculations does highlight the sensitivity of the posterior emissions to background concentration calculations as the comment points out. Therefore, we have changed several parts of the discussions and the conclusions to better reflect the underlying assumptions and uncertainties.

Line 184: Correction factors (1.76) for the AE33 aethalometer at five stations relied on manual verification, with methods for other stations unspecified. This risks introducing systematic biases into observational data.

The manual verification was utilized to make sure that the filter specific correction factor was applied to all sites with AE33 in an effort to remove systematic biases in the entire dataset. This means that our manual verification reduced systematic biases, not the other way around, since cross-instrument comparison showed that AE33 consistently overestimate eBC compared to MAAP. See our answer to Reviewer #4, point 2 for a more detailed description of the process.

Line 200: The manuscript only selected afternoon data (low-altitude sites) or nighttime data (high-altitude sites), discarding >50% of valid observations despite intending to reduce boundary-layer modeling errors. The authors did not verify whether this operation introduces selection bias.

See our answer (3rd paragraph) to Reviewer #4, point 4

From Section 3.2, it can be seen that the measurement of BC utilized multiple instruments and assumed that the MAC values from different sources would not change. However, the AAE of BC is not exactly equal to 1. How did the author consider the impact of these errors on the posterior results?

See our answer (final paragraph) to Reviewer #4, point 4.

Line 261: Remove the redundant parenthesis after “dashed and dotted lines”.

Fixed by removing the parenthesis after “a” on the same line

Line 365: Significant increases in Eastern European emissions during spring/summer (Figure 5) potentially originate from agricultural burning (discussed in Sect. 6), yet no comparison was made with satellite fire detection data or existing fire emission inventories (e.g., GFED) for verification. How did the authors address this issue?

Note that GFED is a part of the non-optimized sector and thus is included in the model. In addition, the main purpose of the manuscript is to test and validate the system. Therefore, we decided to only include a qualitative discussion of the physical mechanisms of the prior-posterior difference. For more info, see our answer to Reviewer #4, point 5.

References

Annadate, S., Arduini, J., Maione, M., Eckhardt, S., Klimont, Z., Stohl, A., Platt, S., 2025. Preliminary emissions of BC for Europe (EU Horizon project PARIS Deliverable No. D7.4). UNIURB, NILU.

Asmi, E., Sipkens, T.A., Saturno, J., Backman, J., Vasilatou, K., Weingartner, E., Keller, A., Ciupek, K., Müller, T., Babu Suja, A., Močnik, G., Drinovec, L., Eleftheriadis, K., Gini, M.I., Nowak, A., Corbin, J.C., 2025. Mass absorption cross-section of ambient black carbon aerosols - a review. Npj Clim. Atmospheric Sci. https://doi.org/10.1038/s41612-025-01288-2

Bergamaschi, P., Segers, A., Brunner, D., Haussaire, J.-M., Henne, S., Ramonet, M., Arnold, T., Biermann, T., Chen, H., Conil, S., Delmotte, M., Forster, G., Frumau, A., Kubistin, D., Lan, X., Leuenberger, M., Lindauer, M., Lopez, M., Manca, G., Müller-Williams, J., O'Doherty, S., Scheeren, B., Steinbacher, M., Trisolino, P., Vítková, G., Yver Kwok, C., 2022. High-resolution inverse modelling of European CH_4 emissions using the novel FLEXPART-COSMO TM5 4DVAR inverse modelling system. Atmospheric Chem. Phys. 22, 13243–13268. https://doi.org/10.5194/acp-22-13243-2022

Evangeliou, N., Platt, S.M., Eckhardt, S., Lund Myhre, C., Laj, P., Alados-Arboledas, L., Backman, J., Brem, B.T., Fiebig, M., Flentje, H., Marinoni, A., Pandolfi, M., Yus-Dìez, J., Prats, N., Putaud, J.P., Sellegri, K., Sorribas, M., Eleftheriadis, K., Vratolis, S., Wiedensohler, A., Stohl, A., 2021. Changes in black carbon emissions over Europe due to COVID-19 lockdowns. Atmospheric Chem. Phys. 21, 2675–2692. https://doi.org/10.5194/acp-21-2675-2021

Evangeliou, N., Thompson, R.L., Eckhardt, S., Stohl, A., 2018. Top-down estimates of black carbon emissions at high latitudes using an atmospheric transport model and a Bayesian inversion framework. Atmospheric Chem. Phys. 18, 15307–15327. https://doi.org/10.5194/acp-18-15307-2018

Grythe, H., Kristiansen, N.I., Groot Zwaaftink, C.D., Eckhardt, S., Ström, J., Tunved, P., Krejci, R., Stohl, A., 2017. A new aerosol wet removal scheme for the Lagrangian particle model FLEXPART v10. Geosci. Model Dev. 10, 1447–1466. https://doi.org/10.5194/gmd-10-1447-2017

Helin, A., Virkkula, A., Backman, J., Pirjola, L., Sippula, O., Aakko-Saksa, P., Väätäinen, S., Mylläri, F., Järvinen, A., Bloss, M., Aurela, M., Jakobi, G., Karjalainen, P., Zimmermann, R., Jokiniemi, J., Saarikoski, S., Tissari, J., Rönkkö, T., Niemi, J.V., Timonen, H., 2021. Variation of Absorption Ångström Exponent in Aerosols From Different Emission Sources. J. Geophys. Res. Atmospheres 126, e2020JD034094. https://doi.org/10.1029/2020JD034094

Jia, M., Evangeliou, N., Eckhardt, S., Huang, X., Gao, J., Ding, A., Stohl, A., 2021. Black Carbon Emission Reduction Due to COVID-19 Lockdown in China. Geophys. Res. Lett. 48, e2021GL093243. https://doi.org/10.1029/2021GL093243

Monteil, G., Scholze, M., 2021. Regional CO_2 inversions with LUMIA, the Lund University Modular Inversion Algorithm, v1.0. Geosci. Model Dev. 14, 3383–3406. https://doi.org/10.5194/gmd-14-3383-2021

Monteil, G., Theanutti Kallingal, J., Scholze, M., 2025. CH_4 emissions from Northern Europe wetlands: compared data assimilation approaches. Atmospheric Chem. Phys. 25, 14251–14277. https://doi.org/10.5194/acp-25-14251-2025

Moosmüller, H., Arnott, W.P., 2009. Particle Optics in the Rayleigh Regime. J. Air Waste Manag. Assoc. 59, 1028–1031. https://doi.org/10.3155/1047-3289.59.9.1028

Müller, T., Fiebig, M., 2021. ACTRIS In Situ Aerosol: Guidelines for Manual QC of AE33 absorption photometer data.

Romshoo, B., Müller, T., Pfeifer, S., Saturno, J., Nowak, A., Ciupek, K., Quincey, P., Wiedensohler, A., 2021. Optical properties of coated black carbon aggregates:

numerical simulations, radiative forcing estimates, and size-resolved parameterization scheme. Atmospheric Chem. Phys. 21, 12989–13010. https://doi.org/10.5194/acp-21-12989-2021
Zanatta, M., Gysel, M., Bukowiecki, N., Müller, T., Weingartner, E., Areskoug, H., Fiebig, M., Yttri, K.E., Mihalopoulos, N., Kouvarakis, G., Beddows, D., Harrison, R.M., Cavalli, F., Putaud, J.P., Spindler, G., Wiedensohler, A., Alastuey, A., Pandolfi, M., Sellegri, K., Swietlicki, E., Jaffrezo, J.L., Baltensperger, U., Laj, P., 2016. A European aerosol phenomenology-5: Climatology of black carbon optical properties at 9 regional background sites across Europe. Atmos. Environ. 145, 346–364. https://doi.org/10.1016/j.atmosenv.2016.09.035