
Reviewer 1

The authors present a significant update to the CAMS-REG inventory, improving the modeling of traffic especially in areas where data are missing. The resulting dataset better aligns with regional inventories and will be of wide interest to the traffic engineering and air quality communities. I recommend publication after the following comments are addressed:

Abstract/introduction:

My understanding is that the authors are updating their CAMS-REG inventory (Kuenen 2022), to a new version (8.1), improving it based on user feedback and new studies. This framing could be clearer in the abstract and introduction, so the reader can rapidly understand the purpose of this paper and dataset. The discussion in section 2.4 is helpful framing and some of that material could be brought into the introduction.

Throughout, but especially introduction: In terms of emissions, your inventory is for NO_x (though of course you generate other useful activity data). This NO_x focus is not clear at the beginning and in the abstract, especially as other pollutants are discussed. Clarify that the particular pollutant that you analyze is NO_x.

Added clarifications that we generate a NO_x inventory in the abstract and in the introduction.
Added clarifications that we base our study on Kuenen et al. 2022 in the abstract and introduction.

Lines 49-73: This literature review is too focused on recapitulating findings from a small number of papers — better to synthesize the general findings of a larger number of studies (e.g. drawing on the findings on papers in 72-73) in a brief paragraph.

Summarised the findings of the recent literature, and included potential reasons for the found underestimation and a small summary.

At the same time, you do not discuss papers using random forest or other ML methods for traffic modeling (cited in section 2.3) which are a significant part of this paper. Consider adding a discussion to the introduction.

Added a section on AADT estimation using machine learning and random forest approaches.

Lines 75, 80: What are these artifacts? How do you know they are artifacts?

Added an explanation of our reasoning for mentioning artifacts.

Line 77: What is a vector scale inventory?

Explained the term.

Figure 1 and introducing text (lines 89-92): It would be useful to say that the Figure shows how the CAMS-REG methodology has been updated/improved since the previous version. Consider defining “links” in the introduction or early in section 2. What precisely makes a road segment a link?

Added a more precise description of the Figure in the introducing text. The definition of road links was added earlier at the enumeration of the goals of the paper.

Section 2.3.1: Define Method 1 and Method 2 in your main text, since you reference it later (in section 3).

Added the model definitions in section 2.3.1.

Line 175: How did you pick 30% and 15% as your fraction for fourth and fifth class road traffic? Is there any way to validate this choice? How sensitive are your NO_x results to a different fraction?

Added sources for reasonable distribution of traffic volumes on smaller roads.

Added Table 'Distribution of absolute and relative shares of NO_x emissions in the domain by OSM class' to give an overview on the importance of smaller roads in respect to the total domain emissions. The roads covered by the spatial gap filling method contribute <4% of total emissions.

Section 3.1: What are the units for AADT? Could you also include a relative error metric (e.g. mean absolute percent error) to help unfamiliar readers understand your MAE/RMSE results? It would also be useful to make very explicit, both here and in section 2, that your goal is to predict AADT as part of your gap filling.

Added an explanation of AADT units in Section 3. Further added clarification on the goal of the gap filling in Section 2.3 and Section 3. We also added MAPE results in the section.

Lines 237-240: The Pulugurtha and Mathew variable importance method used correlations for their linear and GWR approaches. This is less appropriate for random forests, which can make use of nonlinear data (responding differently to different ranges of input data). I would suggest you use a SHAP analysis to investigate variable importance, as it is more robust to the collinearity and nonlinearity of input variables in RFs.

Thank you for catching this. We mistakenly cited the wrong paper here, and changed the citation now to Ishwaran and Lu (2019).

Figures throughout: the text is quite small. Consider increasing font size.

We have updated the figures with smallest text (6,7,8,9) to increase they font size.

Figure 2: Given the volume of the data you plot, consider showing these scatterplots on a log-log scale and/or use binning to transform the points into a gridded heatmap.

Figure 2 was changed to log scale.

Figure 3: Consider adding a panel for changes in urban core emissions for each country, as this general pattern is discussed throughout.

Thank you for this comment. We have changed Figure 3 according to a comment by Reviewer 2 to highlight the changes for each country. In a similar way, we have added a comparison between urban core emissions in the new Figure 4 and accompanying text. For this, we defined urban core emissions as areas covered by the Global Human Settlement Layer.

Figure 4, and the city inventory comparison for London and Milan: A common way to quantify spatial similarity across datasets is the Jaccard index (and there are other similar methods), especially in the GeoML literature. Consider adding a discussion on the Jaccard index to your study when assessing the impacts of your updates on version 4.2 vs 8.1, and on your comparison with the London/Milan datasets. This can be a more robust comparison than spatial correlation alone.

In our understanding, the Jaccard Index is a common way to measure similarity with binary or categorical data. We see no straightforward way to binarize our continuous AADT dataset, and therefore would refrain from applying this index here. As you have pointed out, there are other spatial indices which might offer additional information apart from direct correlation. We added a small section on spatial cross-correlation, as proposed by Chen (2015) (<https://pmc.ncbi.nlm.nih.gov/articles/PMC4438072/pdf/pone.0126158.pdf>) in Section 2.5 (Evaluation methods). We further added spatial cross-correlation results in the Milan and London case studies.

Line 257-258: Do you have a guess as to why your method underestimates emissions compared to regional inventories?

Added a reference to the added results from the AADT comparison.

Section 3.6: Can you go the next step and determine which method (speed-based or land-use-based) is 'better', at least in terms of comparison with regional inventories or other data sets? I want to be convinced that your speed-based method is indeed superior.

Thank you for this comment. Splitting up our emissions into 'urban', 'rural' and 'highway' types is done to apply appropriate emission factors and to be in line with the CAMS convention. When applying speed-based or land-use-based classification of roads into 'urban' and 'rural', we see a different split between those categories, and in the next step also the application of different emission factors. When comparing the 'urban' and 'rural' results between both methods, there can be strong differences, as shown in Figure 9. Summing the results for 'urban', 'rural', 'highway' together, the differences become very small again due to the scaling with country totals and the addition of the different environments, typically less than 1% difference for the cities for which we have regional inventories available. Therefore, these splits do not make a large change for overall totals, and only for 'urban' and 'rural' emissions. We added this result in Section 3.6.

Section 4: I still would like a bit more information on what these artifacts in Eastern Europe are.

We added more information in the Introduction section (around Line 75 of the first submission).

Reviewer 2:

General Feedback

This manuscript presents an original and valuable contribution by improving the spatialisation of NO_x emissions from road transport for the CAMS-REG dataset. The article is well written, logically structured, and the newly generated dataset is both innovative and of high relevance. I would like to congratulate the authors on this achievement.

The authors clearly demonstrated the superiority of the v8.1 dataset relatively to v4.2 across several aspects. These include the higher-resolution spatialisation now enabled by the vector dataset, the notable improvements for non-EU countries, the reduced differences with local city inventories, and the enhanced representation of urban versus rural traffic patterns resulting from speed-based emission factors.

My main concerns relate to the reproducibility and long-term evolution of the methodology. The key input dataset for traffic volumes (OpenTransportMap) was released a decade ago and appears to be no longer maintained. The manuscript does not sufficiently address the temporal representativeness of this dataset, nor does it specify the reference year(s) of OTM. Given that traffic volumes can change significantly over time (particularly in urban areas), and that emission factors evolve rapidly due to the implementation of Euro Standards and fleet renewal, the lack of discussion on temporal aspects is a notable gap. I believe some validation with recent traffic measurements (for instance on the 8 case study cities) could strongly improve the confidence in the resulting dataset.

I have also raised questions regarding the method used to estimate and apply emission factors. The manuscript provides no sufficient information on the distribution of vehicle types across countries or road segments typically. As the data are presented per vehicle category, this omission is important and warrants deeper discussion.

Additionally, I suggested improvements to the readability of several figures, particularly regarding the harmonisation of colour scales and axis ranges. I also recommended some changes and pointed out inconsistencies to improve the readability of the manuscript and the dataset.

Overall, I believe the paper addresses an important topic and has the potential to make a valuable contribution. However, the manuscript currently contains potential errors and unclear explanations that substantially affect readability and interpretation. I therefore recommend major revision. Please accept my apologies for any potential misunderstandings or missed points on my part.

Specific comments

Is this methodology applied to other pollutants in the CAMS-REG-ANT Datasets ? Why was this study focused on NO_x emissions ?

This study focusses on NO_x emissions alone. NO_x is a major pollutant from road transport, and transport NO_x emissions were widely available from other urban inventories for comparison. A second reason why we focus on NO_x here was the compilation of new emission factors, which were completed for NO_x first. In a first estimation, the spatial distribution of NO_x (or the raw spatial distribution of activity data) can also be used for other transport related pollutants. At a later step, we plan to publish spatial distributions for other pollutants based on this activity dataset and newly compiled emission factors.

The manuscript needs clarification regarding the year to which the presented results refer. Specifically, this information is missing or unclear in:

- The captions for Figures 4, 7, 8, and 9.
- Lines 210-211
- Multiple mentions of the AADT data sourced from OTM.
- The dataset

Thank you, we added the years in the Figure captions, line 210 as well as in the abstract and dataset.

Line 7: The "gap-filled activity data" originates from OpenTransportMap, not OpenStreetMap. Consider adding OpenStreetMap on Line 8, following the reference to "... road information data."

Corrected this, and added a reference to the use of OpenStreetMap geometries in Line 12.

Line 7: "n=59,710,490" doesn't match with Table 1 information on road links from OTM data (sum of first column). Could you explain the differences? I would expect the final dataset to have more elements than OTM values since it was not available everywhere?

Thank you for this comment. The number n=59,710,490 wasn't written clear enough in the abstract. This number refers to the number of road links in our final dataset. As we estimated emissions only for the OSM classes listed in Table 3, the number is lower than the complete OTM dataset. In the abstract, we changed 'emissions for all road links' to 'emissions for most road links' to avoid confusion.

Lines 23-24: I usually use EEA numbers (<https://www.eea.europa.eu/en/topics/in-depth/air-pollution/air-pollutant-emissions-data-viewer-1990-2023>) that seem quite different for 2022, do you have an idea to explain these differences? (EEA32-scope): 20.6% for CO, 35.1% for NOx, 15.3% for PM10 and 13.7% for PM2.5

Thank you for pointing this out, changed the values to the ones provided on the EEA data viewer.

Line 96: How is the information on sidewalks used in the manuscript?

Existence of sidewalks was used as one of the predictors in our random forest models. To make the predictors more clear, in Section 2 we added Table 4, which gives an overview on the independent variables used in our models. We also refer to this Table in the Section.

Line 189: Please elaborate on the differences, which are also referenced in Figure 1, between the pure bottom-up approach and the national inventories. Theoretically, these two should align closely, and clarifying the discrepancies would be highly beneficial for bottom-up emission modelers. Specifically, could you quantify the differences (e.g., as percentages) and indicate whether the bottom-up approach results is an underestimation or an overestimation?

We added a more clear description of bottom-up emissions and noted the difference between the bottom-up dataset and the scaled dataset in percent.

Lines 189-190: What is the source of these national totals?

The national totals are reported by countries via UNFCCC or EMEP/CEIP and collected by CAMS. We added a reference to the full documentation (https://atmosphere.copernicus.eu/sites/default/files/publications/CAMS261_2021SC1_D6.1.2-2025_202509_Docu_v2.pdf).

Line 244: Which year are these national total emissions referring to? 2018 (from Figure 3)?

Clarified in the text.

Line 333: “improved NOx emission allocation in 6 out of 7 cities”: I think you used 8 different cities' inventories (Figure 5, line 213) ?

Thank you, we corrected this.

Line 335: typo: “an approximately”

Fixed.

Lines 339-340: I would appreciate further clarification regarding the reported unavailability of OSM data. In my experience, the 'highway' and 'maxspeed' fields are typically well-provided. If data was indeed missing for certain road segments, please describe the methodology used to handle these gaps (only described for missing speed information line 163-164). Finally, I do not understand how the conversion from miles per hour to kilometers per hour could introduce significant errors.

Thank you, we added a reference to the imputation method used for missing data in the independent variables in Section 2.3.1. We further added a short description on the conversion function in Section 2.1.1. The introduction of error refers to the difficulty to catch of labelling conventions entered, for example 'miles/h', 'miles/hour', 'kilometre/h', 'kilometer/hour', and so on. In our case, we only searched for the label 'mph'. This is now also mentioned in the discussion section.

Line 349: For future perspectives on integrating congestion and slower speed effects, maybe you could mention such projects: https://github.com/BlaiseKelly/google_speeds

Added the suggested project together with another research paper utilizing urban congestion values for urban roadside emission.

Lines 358-359: Maybe you could mention this research article, which I believe is closely related to yours: <https://doi.org/10.1016/j.atmosenv.2024.120719>

Thank you for this suggestion, we have added a reference, also at Section 2.3.

Comments regarding the Activity Data and modeling

Line 9: It is difficult to interpret the two different scores (could be misinterpreted as a range) without looking at the methodology as it seems to be for just one “Model”. You could either group the data from Model 1 and 2 all together to present just one number, or you can explain that one score (Model 1) is for EU countries and the other (Model 2) is for non-EU countries with less predictors available.

Thank you, changed.

Line 143: Consider clarifying the term "spatial gap filling" early on, specifying its use only for small roads, as its current appearance is imprecise until later description. Additionally, to avoid confusion, it is recommended to use a more specific name for this approach as your Machine Learning approach is also sometimes referred to as "gap-filling" (e.g., line 110, line 124, line 207).

Added a clarification in the abstract (Line 9). Added a reference to the subchapters on Line 143 to make both methods more clear. As Table 3 highlights, we employ two different methods to generate AADT data, which can be both described as gap filling. To make it more clear which one is used, we now explicitly refer to RF gap filling or spatial gap filling throughout the manuscript.

Lines 154-156: The statement that "information on road class and speed limit was not available from OTM for non-EU Eastern European Countries" appears contradictory, given that you mention having this information, and more detailed, available from the OSM dataset (as stated on line 161 and in Table 3). Furthermore, based on Table 3, it seems OTM road classes were not used (or not useful) for the highway models, as all are categorized as "mainRoad". More generally, the features used in the Random Forest approach (including sidewalks, road surface, two different sources of road classes, land use, and density) are difficult to fully understand. I suggest including a table that summarizes the features employed, their source, and definition (specifying which were excluded for Model 2) to improve clarity.

Thank you, we have added a feature table (now Table 4). We have added a qualification to the statement on Line 161. We added a clarification on the availability of OTM data in Section 2.1.1. Your understanding regarding the highway models is correct. In these models, the OTM road classes are irrelevant.

Lines 173-177: The methodology for estimating low-class road segment traffic volume, despite being challenging, is acceptable as presented and validated against existing literature. However, the use of numerous parameters and associated hypotheses lacks justification. For example, the decision to remove 25% of the data to treat outliers may be excessive. Could you explain the reasons behind selecting these specific parameters? Was the goal, perhaps, to adjust them to achieve an approximate AADT of 50 for these road types?

After a comment by Reviewer #1, we have added references to literature on AADT ratios of smaller road classes. The reason to remove the top 25% of third class road data was due to the sometimes very high values encountered with third-class roads next to highways and the necessity to remove extreme values with one simple method. We shortly expanded on this explanation.

Line 229-230: (related to lines 154-156) Model 2's lower performance (compared to Model 1) may require further discussion: is the absence of certain features the root cause, or are there other factors? This seems counter-intuitive, given the reliance on OSM speed limits and the fact that OTM road class information is captured by the highly influential OSM 'highway' type (line 238).

In the results of our Variable Importance Plots, the maximum speed limits had showed little predictive power on the AADT result. One reason for the loss of model power can be the fuzzy mapping of OTM road class to the OSM highway key (see Table 3), as well as the loss of the other OTM related predictors (Table 4). To derive more clear findings on the interactions of the variables, we would need to perform a deeper analysis into the VIMP calculation.

Line 238: From my understanding, Pulugurtha et al. (2021) applies methods on their coefficients resulting from ordinary least square (OLS) regression and geographically weighted regression (GWR) models. I'm not sure this directly applies to a Random Forest regression, and maybe further explanation is needed to explain how the feature importances are computed in this case.

Thank you, the reference here was wrong. We changed it to Ishwaran and Lu (2019) who provide an overview on Variable Importance calculation.

Lines 354-357: This is perhaps my main concern. I think this point is very important as the OTM dataset was published a long time ago and you don't provide information regarding the temporal

scope for which the dataset was published. Traffic volumes might have experienced substantial changes since this period and especially in urban areas. Moreover, while CAMS has to maintain detailed emission inventories, relying on OTM seems to be a growing challenge in the long term as it appears to have expired (<https://opentransportmap.info>) with their data being made inaccessible (I couldn't access it elsewhere).

We agree, this would be the most important point to solve for future developments, and it would be very beneficial to derive a framework that can combine modeled activity and recent measurement data. Due to the limitations in our project, this was not possible to achieve here.

I strongly agree with the need to obtain up-to-date measurement datasets. You could for instance use this recently published open traffic volume dataset from local measurements on European cities: <https://www.nature.com/articles/s41597-025-05698-y>. I believe this is also a valuable resource as it repertory additional different sources to collect traffic volumes. Could you consider comparing your modeled or OTM-based AADT values with recent measurements from this dataset ? The traffic measurements were map-matched to the OSM network so you might just use the OSMID key directly and I don't think this would be too difficult. I believe this would really strengthen the results from your study if the activity data (which is at the core of the bottom-up inventory and spatial disaggregation) showed an agreement with recent data published by local authorities.

In the long term, you can also consider projects such as AVATAR (<https://avatar.cerema.fr/cartographie>) by the CEREMA in France which gathers near real time and historical traffic volumes on many locations (and not only cities).

Thank you, we added Section 3.2: Validation of activity dataset with measured dataset for the comparison against the suggested paper.

Comments regarding the Emission Factors

Lines 132-133: Figure 1 mentioned that the emission factors are based on the COPERT model, maybe you could mention both sources (both in the figure and in the text) for consistency.

Thank you, we added a reference to the Geilenkirchen et al. (2024) emission factors in Figure 1, which was previously omitted.

Lines 132-141: I feel like this is a really important subsection for the estimation of NO_x emissions but it lacks some details to make it fully understandable and reusable. Please could you provide more details and maybe additional data to clearly present these methodological points ? See some examples below.

Thank you. We have added more details on the intermediate steps of the calculation, and how we use the COPERT data.

Do you have a reference for VERSIT+ data ? If I understood correctly the emission factors (in grams/vkm) for each speed (maybe add this information, as it is mentioned line 147) and vehicle category are from VERSIT+. It would be very useful to provide a dataset for the emission factors used as well (and maybe additional figures, even in supplementary).

Thank you. We have added the relevant references in subsection 2.2 Emission factors.

Then, I'm wondering how the different emission factors (related to vehicle types) are applied to the total traffic volume (which doesn't differentiate between vehicle types) ? Is the distribution of the vehicle types on the roads from COPERT data ?

If this is the case, what is the resolution (in addition to road types) of this data ? Do you have a typical vehicle fleet per country ? Per year ? Do you have different temperature effects on the emission factors for different countries ?

Thank you, we have added these details now. NOx temperature effects are not included in the Dutch national emissions inventory.

Why didn't you use the COPERT emission factor values directly ? It seems highly related since your data is presented with different vehicle types which are similar to the one in COPERT (Buses / HDT / L-cat / LCV / passenger Cars). This could also be presented in the manuscript since the dataset is presented through these categories.

Thank you for the suggestion. While COPERT provides a valuable dataset, we chose to use the Dutch national emission inventory as it is updated annually with emission factors derived from both extensive laboratory and real-world measurements, and serves as the foundation for ongoing work on updated emission factors for different pollutants.

Line 160-161: "The maximum speed of each link was used to decide whether to use urban, rural or highway..." contradicts lines 188 and 308.

Thank you for pointing this out. In Section 2.3.1, we made our method more clear. We first directly assigned the urban class for the smallest roads, and the highway class for the highways. For the rest of the road classes, we applied the method as described.

Line 180: The current description of using "an" emission factor (a single value) is confusing, as multiple values should be necessary to account for the varying "vehicle classes" present on a road segment. Again, I am unclear on the implementation: how are the different shares of vehicles, such as heavy-duty trucks versus light passenger cars, represented for a given road segment?

Thank you, we clarified this at this location. Please also see the changes for the previous comment.

Lines 181: I suggest grouping subsections 2.2 and 2.3.3 as they refer to the Emission Factors calculation and use here.

Thank you, we combined both sections to Section 'Emission Factors'

I am confused by the distinction you draw here between "road class" and "environment (urban, rural or highway)". In section 2.2, you state that "Emission factors... are road-type specific, distinguishing between urban, rural, and highway conditions." Is "road class" distinct from the "environment" types mentioned in Section 2.2, or am I missing something?

Agree, this was confusingly put. Environment (urban, rural, highway) is the sole differentiator between the roads here as well as in COPERT. We changed this in the new combined section.

Line 197: "emission factors are applied": Please elaborate on the application of the emission factors from the previous version. Specifically, were the methodology and the underlying data the same as in previous versions? Providing this detail would help clarify whether the changes observed with v4.2 are solely due to differences in activity data methodology or if an adjustment to the emission factors also contributed to the overall effect.

The emission factors from v4.2 were taken from COPERT. The difference between both versions comes from changes in spatial distribution as well as emission factors. This was updated in the text.

Comments regarding the Figures

Figure 2: The caption says Model 2 was used in EU countries, but the caption of Table 3 and line 230 say otherwise. Please make it consistent.

Changed to make it consistent.

Figure 3: Because country emissions have different absolute ranges, maybe it would be easier to show the evolution (in %) from v4.2 to v8.1 (with +/-%) If you have enough space you could add the new total (in kg/y) of the country on top of it too. I feel like this would help the reader see more clearly what changes occurred on national totals.

We redrew Figure 3 to show relative changes.

For countries which are truncated (like Russia), is it only a subtotal ? How did you scale the initial inventory for these cases ?

Yes, Russia national emissions were only for the European part.

For Figure 3 the unit is kg, but for Figure 5 it's kg/year. I believe kg/year is better.

We gave t/y estimates as Labels in the redrawn Figure 3.

Figure 4: While Panel a) and c) share a common colorscale, Panel b) has a different scale for both versions which prevents the reader from having an unbiased look to these figures. You could consider making one unique colorscale for all the maps here (ranging 0-0.2 for instance). You could consider using the same colormap as in Figure 7 and 8 too since it's the same kind of unit (% of the map total). On the contrary, Figure 9 has a different unit but the same colormap.

Thank you for this suggestion. We agree that readers could not pay attention to the different scales on the left and right plot of Panel b) and might draw wrong conclusions. On the other hand, it is an important message of the picture to show the high concentration of emissions on the left hand plot with its full scale, and not cut off the scale at 0.2 here to align with the right hand plot. We have added a note in the Figure description to pay attention to the different labels in Panel b).

Figure 5: The two panels have the same y-axis label but different ranges, which may confuse the reader: please standardize the range. The different ranges appear to be necessary due to large emissions, particularly from London and Paris. In the case of Paris, the high values are mainly attributable to the decision (line 216) to include the entire Île-de-France region, which encompasses numerous rural and low-density areas. Note that the AirParif inventory is also available specifically for the Paris municipality. Additionally, the current link to the inventory is expired, please consider using this one:

<https://www.airparif.fr/surveiller-la-pollution/les-emissions#> (3187 tNO_x for the year 2018).

London is known to have an extensive area (contrary to Paris) so this might be less surprising (although Birmingham seems comparable).

Moreover, from the "cities_polygons.gpkg" file, we can observe that the Barcelona geometry seems to be overly simplistic: is it comparable to the local inventory ? Other geometries used also show simple square-like shapes, and I'm wondering how this is comparable to the others (Figure

5). For instance, part of the Munich or Krakow ring-road is not included in the geometry. The utilization of a standardized urban area boundary dataset is suggested as a more straightforward and academically correct approach.

Thank you for this comment. For the cities which are compared to local inventories (e.g. Barcelona), we have created the geometry in `cities_polygons.gpkg` to follow the local inventory which was provided to us. The boundaries of the other cities were created by us. To make these more in line with a standard dataset, we have replaced these by the boundary shapes provided here: <https://ec.europa.eu/eurostat/web/gisco/geodata/statistical-units/cities-functional-urban-areas>. Please note that these are only available for cities within the EU. For non-EU cities, we have kept our previous boundaries. We have updated the `cities_polygons.gpkg` and uploaded them to a new version of our files under the same DOI.

Figure 7 and 8: To facilitate comparison between the maps, please ensure a consistent color scale is used, ideally combining them into a single colorbar.

Thank you, we have fixed the colourbars.

Comments regarding the Dataset

The `readme.txt` file should contain more information regarding the naming of gridded files to simplify usage, as there seem to be 4 dimensions:

- road type: all / highway / urban / rural
- variable : emission : activity (vkm)
- vehicle types (COPERT types) + “sum”
- model: corine or speed-based

Thank you, added.

In the filenames: “commerical” → “commercial”

Replaced.

The two datasets (gridded and vector) are in different CRS which can lead to confusion.

Thanks, we saved the vector dataset to EPSG:4326.

For the gridded datasets, looking at the sum of emissions (6237 tNO_x), the unit of the provided emission files appears to be inconsistent with the previously cited EEA numbers (2.5 - 3 MtNO_x) and the CAMS-REG v8.1 dataset from ECCAD (2.84 MtNO_x). A discrepancy factor of approximately 455 might suggest that the gridded emissions might be expressed in kg NO_x per day, which aligns with the daily temporal resolution of the 'vkm' variable in the vector dataset. Similarly, the unit for "vkm" files needs clarification. Given the resulting value when summing everything (7.7 Million), it is unlikely that the result is in "daily vkm". For instance, Paris has a yearly traffic of approximately 5 Gvkm (Airparif). Maybe it is thousands of vkm ?

Thank you a lot for this valuable comment. Looking over the units again, in the `.gpkg` file vkm is indeed in 10^3 / day, and emission is in g/day. The gridded files are in kg/day. We have changed this in the `Readme.txt`.

The large size of the vector dataset (over 40GB in a single file) may discourage potential users. Could the same segmentation approach used for the gridded file (96 files) be applied to the vector dataset to produce smaller, more manageable files (the current file has 48 columns)? Additionally,

it would be beneficial to include the retained maximum speed for road segments in the data (see below).

Thank you for this suggestion. We have added the maximum speed columns in the dataset. While it is true that the final dataset is very large, we feel that a single .gpkg file is still the most handy way to store this dataset. Splitting the data columns into separate files, while needing to retain the geometry and some base information for each road, would still require ~20GB for each file and create large duplications, and good tooling for subsetting .gpkg files. We have added a note in the Readme.txt to be aware of the large file size.

Applying the city boundaries masks to the dataset I found significant differences between Figure 5 and your dataset for total yearly emissions in some cities, notably Moscow (2000 tNOx), St Petersburg (1349 tNOx), or Istanbul (8873 tNOx). Please verify the numbers are correct, I provided some code attached to explain how I obtained these numbers.

Thank you for noting this. Our old results were based on an outdated version of the gridded results. It is certainly more precise to base the city comparison on the vector dataset as performed in your code. We have applied our new cities_polygons.gpkg and recalculated the city emissions accordingly and changed the accompanying text and the Figure, as well as shares of urban emissions and the accompanying Figure.

In an attempt to look at the emission factor values you used, I divided the (unscaled) emissions by the (converted to yearly basis) vkm columns. It appears the emission factors (in grams/vkm) are pretty high, notably for Passenger Cars as they exceed 1 gNOx/vkm, although most cars in the EU now have Euro Standards and should have lower emissions associated. Here the retained speed could also help the user understand these values. Please comment and add more detail on this, also for other vehicle categories.

Thank you for this comment.

For France, the urban emission factors used by us are:

0	Buses	4.27350832
1	Heavy Duty Trucks	5.14658474
2	L-Category	0.05825076
3	Light Commercial Vehicles	0.86124429
4	Passenger Cars	0.48241182

These are weighted by the vkm shares (in percent). As the vkm are from Copert, we can not share the data, but passenger cars have around 75% share.

We tested the ratio between emission and vkm, and came to 0.672 g/vkm for Paris in our dataset.

This double analysis (vkm and emission factor) really helps to better understand what are the key differences between different modeling approaches. For instance, in Figure 5 one could say that you have close agreement with AirParif regarding the total emission in Île-de-France on 2019, but a closer look shows significant differences between the parameters (p.51-54):

https://www.airparif.fr/sites/default/files/document_publication/Bilan_IDF_2022_0.pdf

	NOx emissions (tons)	Activity (Gvkm)	Implied EF (gNOX/vkm)
This study	32594	25	1.30
AirParif	31720	~52	0.61

Additional changes:

Removed 'trunkline' from Table 3, as this OSM class was not used in the final dataset.