Link-based European road transport emissions for CAMS-REG v8.1 and a comparison to city inventories

Tilman Leo Hohenberger[1], Marya el Malki[1], Antoon Visschedijk[1], Marc Guevara[2], Martin Otto Paul Ramacher[3], Alessandro Marongiu[4], Guido Giuseppe Lanzani[4], Giuseppe Fossati[4], Anu Kousa[5], Eleni Athanasopoulou[6], Anastasia Kakouri[6], and Jeroen Kuenen[1]

[1]Air Quality and Emissions Research, TNO, Princetonlaan 6, 3584 CB Utrecht, the Netherlands
[2]Barcelona Supercomputing Center, Barcelona, 08034, Spain
[3]Helmholtz-Zentrum Hereon, Max-Planck-Str. 1, 21502 Geesthacht, Germany
[4]ARPA Lombardia, Environmental Protection Agency of Lombardia Region, 20124 Milano, Italy
[5]Helsinki Region Environmental Services Authority, Ilmalantori 1, 00240 Helsinki, Finland
[6]National Observatory of Athens, Vas. Pavlou & I. Metaxa, Penteli GR-15 236, Greece

Correspondence: Jeroen Kuenen (jeroen.kuenen@tno.nl)

Abstract. Spatially resolved estimates of road transport emissions are fundamental for tackling challenges of air pollution and greenhouse gas emissions. Emission estimates at 0.05° x 0.1° resolution are provided in the widely used CAMS-REG regional European emissions inventory. ~~For~~ Building on previous work by Kuenen et al. (2022)) or the road transport sector, several improvement opportunities were identified: Firstly (1) an underestimation of ca. 35% of NOx emissions in comparison to 8 independent urban inventories; secondly (2), artefacts in the spatial distribution in Eastern European non-EU countries; thirdly (3), the necessity of labour-intense downscaling methodologies to create high-resolution urban inventories from the fixed resolution dataset. To overcome these, emissions for ~~all~~ most road links in the domain (n=59,710,490) were estimated using gap-filled activity data (AADT) from ~~OpenStreetMap and OpenTransportMap.~~ OpenTransportMap, targetting NOx for the base year 2018. Gap filling was performed with random forest models trained on land-use and road information data ~~. Model~~ and with a spatial method for small roads. For non-EU countries, model performance was R^2:0.63~~-0.74 and~~ , MAE(AADT):~~1570-2028~~2028, for EU countries it was R^2:0.74, MAE(AADT):1570, with better performance for larger roads. Up-to-date emission factors for NOx were applied on road links using a novel maximum-speed–based classification. To generate the CAMS-REG v8.1 inventory, the resulting spatial distribution was used as a proxy map, together with national totals.

The new dataset ~~lowered the difference-to-city inventories to 19~~is based on OpenStreetMap geometries and lowered the difference to city inventories to 18% for absolute NOx emissions~~, and~~ . It can be flexibly gridded to high resolutions. ~~Median increase in urban emission share is 24% compared to national totals, and~~ Some non-EU cities see large increases (~~e.g. Istanbul~~Kiev, +~~197%; St. Petersburg~~84%, Istanbul, +~~288~~360%) in attributed emissions due to the updated spatial distribution. Two case studies (London and Milan) show an increased spatial correlation, from $R^2 \approx 0.3$ using CAMS-REG v4.2 to $R^2 \approx 0.6$, with CAMS-REG v8.1 against the local inventory. Vector and gridded versions of the emission dataset and spatial distribution are available at https://doi.org/10.5281/zenodo.15688723 (Hohenberger et al. (2025)).

Emission inventories are a fundamental starting point for modelling and policy analysis for both greenhouse gases (GNG) and air pollutants. Coupled with air quality models, they are used to derive pollutant concentrations (El-Harbawi (2013)), epidemiological studies (World Health Organization and others (2021)), exposure analysis (Isakov et al. (2009)), source apportionment (Hopke et al. (2022)) or sectoral analysis (Gu et al. (2018)). Road transport is a major source of air pollution and GHG emissions. Shares of Europe-wide emissions were as high as ~~15.6~~20.6% for CO, ~~26.6~~35.1% for NOx, ~~8.5~~15.3% for PM_{10} and ~~7.6~~13.7% for $PM_{2.5}$ ~~,~~in 2022 (~~?~~EMEP Centre of Inventories and Predictions (2023)). Road transport plays an important role in exposure to air pollutants, especially in cities (Nieuwenhuijsen (2016); Gurram et al. (2019)), where children and households in poverty are often overly affected from exposure to road-based air pollution (Barnes et al. (2019)). Contributions to $PM_{2.5}$ emissions in urban areas throughout 150 European cities are on average 15% (Thunis et al. (2023)). Increasing the spatial resolution of emission inventories beyond the typical kilometre-scale is important especially for urban applications. The use of low-resolution emission data in urban areas has been found to underestimate health effects (Gurram et al. (2019)). In cities, larger population density meets relatively high emissions, for example from road transport, and often complex urban terrains, which can result in steep concentration gradients (Santiago et al. (2021); Fu et al. (2020)). Together, these features require higher resolution emission inventories as a basis for different modelling techniques (Kadaverugu et al. (2019); Yang et al. (2019)).

In Europe, city-scale emission inventories are available for a number of major cities, including London, Paris, Barcelona and Helsinki (see 2.1.4). Here, the cost, effort and knowledge needed for inventory preparation can be a barrier for administrations, which limits urban modelling and planning efforts at this level.

In an attempt to increase the spatial resolution, it is possible to downscale existing gridded national, regional or global inventories. This downscaling is performed using proxies such as local road geometries, road type or activity data (Trombetti et al. (2017); Valencia et al. (2022). Manual downscaling using proxy data is again labour intensive, and requires careful data handling and result verification. This can be mitigated by tools such as *UrbEm*, which is available for European cities to downscale regional emissions to city scale (Ramacher et al. (2021)).

Combined with emission factors and fleet composition information, a complete set of roadside activity data allows the calculation of road-level emissions, thus circumventing the need for re-gridding or downscaling of a fixed-size inventory. Roadside activity data is often given in the form of AADT (Annual Average Daily Traffic). However, available activity data for roads is often incomplete. Where ever necessary, gap filling activity data is therefore an important exercise for calculating complete roadside emission. Methods for predicting missing AADT values include statistical modelling and machine learning methods. Recent studies could show improved results with neural networks and tree-based methods compared to statistical approaches Han (2024); Sekuła et al. (2018). Amongst the machine learning approaches used to estimate AADT data, random forest (RF) is a common choice among transport modellers (Wen et al. (2022); Baffoe-Twum et al. (2023); Das and Tsapakis (2020); Sekuł). Random forest models are flexible when working with missing and incomplete data (Tang and Ishwaran (2017)).

The European Copernicus Atmosphere Service (CAMS) maintains a number of regional and global air quality and GHG emission inventories for anthropogenic and biogenic emissions (Denier van der Gon et al. (2023)). Amongst them, CAMS-REG-ANT (hereafter CAMS-REG) is an emission dataset covering UNECE-Europe at 0.05° x 0.1° resolution, and is used in the CAMS Regional Air Quality Production System (Colette et al. (2024)), as well as in a wide number of studies on air quality and public health (recent studies include: Khomenko et al. (2023); Achebak et al. (2024); Chang et al. (2024)). In CAMS-REG, reported country emission sector totals are distributed spatially via gridded proxy maps on a national level, without any prior intermediate regionalisation steps (e.g. to provinces or municipalities). A detailed overview on CAMS-REG is available at (Kuenen et al. (2022)).

~~Common feedback from CAMS-REG users has been the underestimation of urban road transport emissions. The spatial road transport distribution for~~ The spatial distribution of road emissions in CAMS-REG has not been updated since ~~v4. The underestimation is presumably caused due to a proportionally small share of national emission totals being distributed to urban cores. Several studies reveal this underestimation.~~ this version 4.

~~Skoulidou et al. (2021) validated Chemical Transport Model (CTM) simulations of NO2, based on CAMS-REG emissions, for Athens and Thessaloniki, with 14 measurement stations as well as with multi-axis differential optical absorption spectroscopy (MAX-DOAS). In both cases,~~ This study aims to update on the approach and dataset described by Kuenen et al. (2022), focussing on NOx. There is a growing pool of studies aiming at verifying modelled transport emissions with measured air pollution concentration in European cities. The degree of underestimation is highly variable throughout cities and methodologies. In the following, an overview on recent studies and a subsequent summary of the most important findings is given.

Reported underestimation of modelled ~~urban NO2 concentration were found (mean bias~~ concentration is common throughout cities in Europe. Skoulidou et al. (2021) report the degree of underestimation at traffic sites for NOx in Athens at -10% (mean bias) against measurement stations; and at -2 to -35% against spectroscopy based MAX-DOAS ~~).~~

~~Another study performed by Navarro-Barboza et al. (2024) aimed at modelling carbonaceous aerosols within Spain. Model simulations were based on CAMS-REG v4.2 and the HERMES v3 emission model, which includes link-level on-road emissions for Spanish roads (Guevara et al. (2020)). Whereas total emissions for PM are consistent for both emission models, the spatial and sectoral aggregations reveal large differences. In Barcelona , CAMS-REG assigns ~10% of total~~ measurements. In Hamburg, the degree of underestimation for NO2 at four traffic sites was given at 20% (Karl et al. (2024)). Pültz et al. (2025) further find a systematic underestimation throughout urban background NO2 sites in Germany. Kuik et al. (2018) finds an underestimation of urban NO2 of 30% at the urban background and of 50% at the urban core in the Berlin metropolitan area. A study focussing on Barcelona finds a much lower attribution of PM2.5 ~~emission~~ and BC emissions to the transport sector ~~, contrasting with HERMES ' ~~~~in CAMS-REG, compared to the HERMES emission model (10% in CAMS-REG compared to 70% ~~. Similar discrepancies were found for BC, where a CAMS-REG–based simulation greatly underestimated the traffic contribution at an urban background site in Barcelona (18.5-25.8% modelled traffic contribution against 78.8-84.2%measured fossil contribution)~~in HERMES) (Navarro-Barboza et al. (2024)). Underestimation of transport emissions are further mentioned for Marseille (Karl et al. (2023)) and Liège (Ramacher et al. (2025). Throughout these recent studies, the underestimation of urban transport emissions for NOx in European cities was thus found to range between 10-50%.

~~During the validation of their air quality forecasting system urbanAQF in Hamburg, Karl et al. (2024) performed a year-long validation of 2021 forecasts against measurement stations, with Emissions based on~~ Out of the presented works, Skoulidou et al. (2021), Navarro-Barboza et al. (2024), Karl et al. (2024) directly use the CAMS-REG ~~v4.2. Transport emissions were downscaled with the UrbEm tool (Ramacher et al. (2021)), and further allocated to road links using OSM road classes. Lastly, road emissions in the urban centre were multiplied by a factor of three to account for potential underestimation (based on results from~~ inventory, and are thus highly relevant for this study. The spatial distribution is given as a common reason for the found underestimation (see Karl et al. (2023), Pültz et al. (2025), Kuik et al. (2018)). ~~Whereas the system was able to reproduce measured NO2 concentrations at urban background and industrial sites, modelled concentrations at the four traffic sites saw a 20% underestimation of concentrations. Contrasting this, PM2.5 concentrations were overestimated at traffic sites. The authors hypothesize that this was caused by the use of an automatic correction method based on PM10 observations. In contrast to PM2.5, PM10 concentrations are more stronlgy influenced by non-exhaust particle resuspension (Karl et al. (2024)).~~

~~Underestimation of transport emissions are further mentioned for Marseille (Karl et al. (2023)), Liège (Ramacher et al. (2025) and in German cities (Pültz et al. (2025))~~Apart from the spatial distribution, further other reasons are presented. In their model setup, Kuik et al. (2018) discuss boundary layer height and model mixing as potentially impacting their results. Moreover, Kuik et al. (2018) discuss the uncertainty of activity data and emission factors, and Pültz et al. (2025) highlight the potential impact of a misallocation of cold starts away from the urban core.

Lastly, the spatial distribution in previous CAMS-REG versions in non-EU Eastern European countries such as Ukraine or Russia showed ~~many artifacts , and city emissions~~ a number of artifacts (see Figure 5), where a large portion of the country's emission is attributed to a single pixel or road. For Ukraine, the highway E105 between Dnipro and Kharkiv received the largest share of national emissions, with emission values far exceeding the largest cities in the country. For Russia, a similar distribution was noticed between the highway M-11 between Moscow and St. Petersburg. Moreover, city emissions for non-EU large cities were not on levels comparable with other European cities (see Chapter 3.4~~and Figure 5).~~ . These artifacts are linked to the low level of road information available at the time of preparation and the activity data estimation method described in Kuenen et al. (2022).

Our update to the road transport emission distribution therefore targetted:

1. ~~a vector scale inventory~~ an inventory on the level of OSM road segments (road vectors, in the following also called road links) that can be flexibly scaled to high resolution and urban scenarios;
2. a distribution of more emissions to urban areas, in line with previous comparisons based on city inventories and previous studies; and
3. a reduction of artefacts, especially in non-EU Eastern Europe.

The new distribution has been included in CAMS-REG since v8, with v8.1 being the latest published version. While staying close to the datasets used in previous CAMS-REG versions for consistency, we attempted to update the spatial distribution based on ~~holistic~~ gap filling of activity data and urban/rural status for all road classes. We further updated the emission factor

set, and produced vector-based emissions for the majority of roads within our domain. These were then gridded to city scale, and compared with the best available urban inventories, which is further explained in the methodology (Section 2). Section 3 presents the results of our model, and shows comparisons with a number of urban inventories and the spatial distributions of previous CAMS-REG versions. The discussion can be found in Section 4, and data availability is described in Section 5.

2 Methodology

~~An overview of the methodology is given in Figure 1~~Figure 1 gives an overview on the methodology changes between CAMS-REG 4.2 and CAMS-REG 8.1. The following subsections describe the components of the methodology, data, emission factors, statistical methods and evaluation criteria in detail. Chapter 2.4 compares the methodology to the previous methodology of the spatial road transport distribution used in CAMS-REG versions 4.2-8.1. All updates presented in this work are included in CAMS-REG since v8.0.

Figure 1. Overview of CAMS-REG-4.2 and CAMS-REG-8.1 methodologies for road transport.

2.1 Data used

2.1.1 Transport data

The road vector dataset used in this study was retrieved from OpenStreetMap (OpenStreetMap contributors (2023)). Information on sidewalks, number of lanes, one-way street status, road type and speed limit was extracted from the OSM dataset with the help of parsing functions wherever available. Speed limits in ~~miles~~mp/h were transformed to km/h to achieve a consistent dataset using a function that scans the speed limit entry for 'mph' and applies the conversion factor.

We used the OpenTransportMap dataset (Jedlicka et al. (2016)) for link-level activity in the form of ~~an Annual Average Daily Traffic (AADT) value~~AADT. OTM divides roads into six categories ranging between main roads and fifth-class roads. A detailed methodology is available in Jedlicka et al. (2016). Additionally, information on the surface of the road, speed limit and the road capacity is available. The dataset offers substantial information on road activity, especially for major roads. However, there are large data gaps especially for smaller roads, and no AADT values are provided for the smallest road classes, while the majority of individual links in the fourth or fifth class (Table 1). The OTM dataset was then combined with OSM geometries and data columns using a shared reference column. OTM data is available for all 27 EU countries, as well as the UK, Norway and Switzerland. For non-EU countries in the Eastern European region but included in the CAMS-REG domain (e.g. Ukraine, Serbia, western Russia), traffic data from OTM was not available.

Road class	n	AADT mean	AADT SD	%Available
Main Road	1,124,835	8,826	12,274	78.9
1st class	2,138,322	4,834	7,249	85.3
2nd class	3,672,761	1,708	3,519	79.1
3rd class	5,714,366	581	1,920	54.8
4th class	27,845,043	-	-	0
5th class	34,601,623	-	-	0

Table 1. Data availability and general statistics on OTM data. %Available shows the share of non-zero AADT values for all roads. AADT SD shows the Standard Deviation of AADT values within a road class.

2.1.2 Countries in the study area

The study area covers UNECE-Europe, including the whole of Turkey and the European part of Russia (up to 60°E). Countries in the study area, but without available OTM data are Albania, Bosnia, Serbia, Moldova, Belarus, Ukraine, Russia, Georgia, Azerbaijan, Armenia and Cyprus. For these countries, spatial distributions were derived fully using ~~gap-filling~~ gap filling models (see Section 2.3).

2.1.3 Spatial data

We obtained the CORINE land cover dataset (2018 version) from Copernicus Land Monitoring. This dataset gives high-resolution information on land use in a 100x100 meter raster grid, and 44 categories for 39 European countries, but excluding large countries within our domain such as Russia, Ukraine and Belarus (Copernicus (2018)). The land-use classes are roughly divided into artificial surfaces, agriculture and forest/natural areas. The data was kept in its initial resolution.

To obtain land use information for the part of the study domain where CORINE was not available, we further used the ESA Worldcover dataset (Zanaga et al. (2022)). ESA Worldcover consists of 11 land use classes with global coverage at a 10-meter resolution.

For population density data, we queried the LandScan Global dataset for the year 2021 (Bhaduri et al. (2002)). The dataset is in a resolution of 30 arcseconds (approximately 1 km), and gives values for urban and rural population within each grid cell. Landscan data was available for the entire study domain.

We then extracted spatial information from the aforementioned mentioned raster datasets for each OSM link, which was then used for training and predicting of ~~gap-filling~~ gap filling models to generate missing traffic density values (see Table 3). Additionally, country information was added with the help of a country mask.

2.1.4 City road traffic emission inventories

We used city-scale urban inventories independently generated by local or national authorities to compare our emission estimations on the local level, with the assumption that urban inventories incorporate the most detailed local information. Table 2 gives an overview of city scale inventories that were made available to us, for the sake of this comparison, by local authorities or institutes.

City	Year	Reference
Barcelona	2018	Guevara et al. (2020)
Birmingham	2018	National Atmospheric Emissions Inventory (2025)
Bologna	2017	Arpa Emilia-Romagna (2025)
Hamburg	2018	Samland et al. (2024)
Helsinki	2019	Finnish Meteorological Institute (2025)
London	2019	Greater London Authority (2023)
Milan	2019	Marongiu et al. (2024)
Paris	2019	Airparif (2025)

Table 2. Summary on local NOx inventories for road traffic used in this study

2.2 Emission factors

For the calculation of NOx emissions from road transport, this study used emission factors based on the Dutch national emissions inventory; for details and the full dataset, see (Geilenkirchen et al. (2024)). The methodology is vehicle-specific and uses a bottom-up approach introduced in 2019, which calculates emissions at the level of individual vehicles using annual odometer readings from the RDW (Netherlands Vehicle Authority). Each vehicle is assigned to one of approximately 350 VERSIT+ vehicle classes, defined by vehicle type, weight, fuel type, emission legislation category (i.e., Euro standards), and exhaust after-treatment technologies Smit et al. (2007); Grange et al. (2019). Emission factors for NOx are expressed in grams per vehicle kilometre and are road-type specific, distinguishing between urban, rural, and highway conditions. These emission factors are derived annually from both laboratory testing and real-world driving measurements, and are adjusted for vehicle aging using odometer data. ~~The final NOx emissions are calculated by multiplying these factors by annual kilometres driven per road type based on COPERTdata (Ntziachristos et al. (2009)~~NOx temperature effects are not included in the Dutch national emissions inventory. Applying vehicle kilometres and vehicle type information from COPERT (Ntziachristos et al. (2009)) to our emission factors dataset, we calculated the national emission totals for each country. In a next step, we derived weighted average emission factors per sector and vehicle kilometre. For this, the national emissions were divided by the total sector vehicle kilometres from COPERT. The COPERT dataset supplies yearly annual data on national vehicle fleet composition, and can readily be used to divide the dataset into relative shares for each vehicle category.

For each road link, the corresponding weighted emission factor was multiplied with the road's AADT value and road length. The weighted emission factor is dependent on the location (country) of the road, as well as the environment (urban, rural, highway) and vehicle class (see Section 2.2), resulting in ~~detailed and representative estimates of NOx emissions from the European vehicle fleet~~an annual emission per year and vehicle class. In previous works, the urban–rural split was performed with gridded land use or population density data, resulting in all roads in an area sharing the same emission factors. However, emission factors of roads depend on the road conditions, maximum speed and congestion levels (Int Panis et al. (2006); Li et al. (2016); Fra), and not on the location of a road within an urban boundary. In an effort to choose the appropriate emission factor on a link level, we used urban emission factors for all roads with a speed limit of $\leq$50km/h, and rural emission factors for all roads with a speed limit of >50km/h. Maximum speed information was preferentially taken from OSM where available, or from OTM as a fallback. Highways were identified by road class and received their own emission factor independent of their speed limit.

The resulting dataset is referred to in Figure 1 as *bottom-up* emissions. The term bottom-up emissions stands for the dataset which was based directly on our activity dataset. We then scaled bottom-up emissions with reported national totals (for more details, see Denier van der Gon et al. (2023)) and gridded with 0.05° x 0.1° resolution to generate our final results for CAMS-REG v8.1 in line with the reported data. Both datasets aligned closely, both bottom-up emissions overestimating domain totals by 3%.

2.3 Statistical methods

We used a combination of random forest (RF) models (see Section 2.3.1 and spatial gap filling (see Section 2.3.2) to create a complete dataset of AADT and maximum-speed values for most roads in the study area. This is done to generate a complete activity dataset. On this basis, the road-level emissions are then calculated in a next step. Gap filling AADT data with machine learning approaches (Han (2024); Sekuła et al. (2018); Razali et al. (2021)) among them, random forest (~~Wen et al. (2022); Baffoe-Twum et~~ Wen et al. (2022); Baffoe-Twum et al. (2023); Das and Tsapakis (2020); Sekuła et al. (2018); Shen et al. (2024)) is a common choice among traffic modellers, especially when working with mixed and incomplete data. AADT values were later used to calculate absolute NOx emissions per road. A complete set of speed limits is necessary to choose the appropriate emission factor per road.

2.3.1 Random forest models

We first matched OSM road classes to their corresponding OTM class (Table 3). OSM links with road classes other than those given in Table 3 were ignored. Subsequently, RF models were trained to predict AADT values for OSM links without available AADT data from OTM. After initial testing, it was noticed that a single RF model over all modelled road classes resulted in larger model errors for high-intensity highways; therefore, we trained an additional highway model, predicting only for missing highway links (Table 3). Information on road class and speed limits from the OTM dataset, as well as land use data from CORINE, was not available for the previously outlined non-EU Eastern European countries. For the use in these locations, we trained RF models that excluded these variables. In total, we thus trained four different RF models, differentiating between highway/non-highway as well as OTM data availability. These models are called RF model 1 (for non-highway roads in EU countries), RF model 2 (fore non-highway roads in non-EU countries), RF highway model 1 (for highways in EU countries) and RF highway model 2 (for highways in non-EU countries). An overview is given in Table 3. An overview on the independent variables used in both model types is givein in Table 4.

For RF modelling, we used the randomForestSRC library (version 3.3.1) (Ishwaran and Kogalur (2024)). The number of trees was set to 250 trees per model, with 10% of data reserved for testing. All other settings were kept to the library defaults. Missing predictor data was handled by imputation with random forest (for details see Ishwaran and Kogalur (2024)).

The maximum speed of each link was used to decide whether to use urban (<50km/h), rural (50km/h -100km/h) or highway emission factors (>100km/h) ~~. Information~~ in the case were a straightforward identification by road class was not possible. We first categorized all fourth- and fifth roads key as urban, and all mainRoad, motorway and highways (OTM classes) as highway. For other road classes, information on maximum speed was taken from OSM and OTM data where available, with a preference for OSM data, as this information came from a more recent dataset. Then, all maximum speed data were transferred into the three outlined categories. Subsequently, one RF model with the above mentioned settings was trained to predict these categories for all links missing maximum speed information.

OSM class	OTM class	~~Gapfilling~~ Gap filling method
trunk	mainRoad	RF highway model 1 / RF highway mode
motorway	mainRoad	RF highway model 1 / RF highway mode
motorwaylink	mainRoad	RF highway model 1 / RF highway mode
~~trunkline mainRoad RF highway model 1 / RF highway model 2~~ primarylink	firstClass	RF model 1 / RF model 2
primary	firstClass	RF model 1 / RF model 2
secondary	secondClass	RF model 1 / RF model 2
tertiary	thirdClass	RF model 1 / RF model 2
tertiarylink	thirdClass	RF model 1 / RF model 2
residential	fourthClass	Spatial gap filling
livingstreet	fourthClass	Spatial gap filling
track	fifthClass	Spatial gap filling
service	fifthClass	Spatial gap filling

Table 3. Matching table between OSM and OTM classes and gap filling approach used. Model 1: Model includes OTM predictors. Model 2: Model excludes OTM predictors, for use in non-EU countries

Independent variable	Unit	Model Type 1	Model Type 2
OSM Highway class	Categorical	X	X
Landuse class	Categorical	X	X
Road Length	Continuous (m)	X	X
Landscan Urban Population	Continuous (number of people)	X	X
Landscan Rural Population	Continuous (number of people)	X	X
Landscan Total Population	Continuous (number of people)	X	X
OSM Sidewalk	Categorical	X	X
OSM Lanes	Continues (n)	X	X
OSM Road Surface	Categorical	X	X
OSM Maximum Speed	Continuous (km/h)	X	X
NUT 3 Code	Categorical	X	-
OTM Highway class	Categorical	X	-
OTM Road Surface	Categorical	X	-
OTM Form of Way	Categorical	X	-
OTM Road Direction	Categorical	X	-

Table 4. Overview on predictors used for Model Types 1 and 2. Model 1: Model includes OTM predictors. Model 2: Model excludes OTM predictors, for use in non-EU countries

2.3.2 Spatial gap filling

For local roads (forth and fifth class roads), no information on traffic volume was available from OTM. In our dataset, these local roads account for 79.1% of the total road length, and it was thus seen as important to also estimate AADT values for all local roads, even if the traffic volumes were expected to be low for each road. For these roads, we used a gap filling approach based on the average traffic volume of tertiary roads in the surrounding area (see Table 3). Approaches for estimating local road data in the literature include regression approaches (Apronti et al. (2016); Pulugurtha and Mathew (2021); Sun et al. (2015)), spatial techniques such as kriging or inverse distance weighting (Baffoe-Twum et al. (2023)) or machine learning approaches (Baffoe-Twum et al. (2023); Sharma et al. (2001)). After performing gap filling for larger roads using our RF method (Section 2.3.1), we divided the study area into 5x5km grids, and removed the top 25% of tertiary roads with the highest traffic volume within each grid cell to handle outliers~~, such as misclassified roads or roads that were connected to highway exits~~. Removal of the top quantile was done due to a number of third class roads carrying a very high traffic load (for example highway exits), which were deemed less relevant for the prediction of residential roads. Then, we took the mean of the remaining tertiary roads' traffic volume within each cell. The traffic volume of fourth- and fifth class roads were set to 30% and 15% of that value, respectively, resulting in a minimum AADT of 15 for fourth class roads, and a minimum AADT of 8 for fifth class roads. Macbeth and Ltd (2007) give the ratio of typical values between urban collector (which maps to our third class roads) and local roads to around 30%. Cornell Local Roads Program (2014) give a typical cut off between low- and high volume roads at an AADT of 1000. Further, the cut off between low- and very-low volume roads is given at around 400. Our judgement on reasonable percentage-base values for forth- and fifth class roads was made based on these sources.

The resulting raster values used to estimate small roads' AADT ranged from 50 to 3358, with a mean of 57. This results are comparable with estimations from previous studies (Pulugurtha and Mathew (2021)). The smaller roads on which spatial gap filling was performed result to less than 4% of total domain emissions (see Table 1). Therefore, the spatial gap filling method does not have a large effect on domain totals.

2.3.3 ~~Use of emission factors, scaling and gridding~~

~~In a next step, we added our previously compiled emission factors to the traffic intensity values. For each road link, an emission factor was chosen based on country, vehicle class, road class and environment (urban, rural or highway), resulting in an annual emission per year and vehicle class. In previous works, the urban–rural split was performed with gridded land use or population density data, resulting in all roads in an area sharing the same emission factors. However, emission factors of roads depend on the road conditions, maximum speed and congestion levels (Int Panis et al. (2006); Li et al. (2016); Franco et al. (2013)), and not on the location of a road within an urban boundary. In an effort to choose the appropriate emission factor on a link level, we used urban emission factors for all roads with a speed limit of $\leq$50km/h, and rural emission factors for all roads with a speed limit of >50km/h. Maximum speed information was preferentially taken from OSM where available, or from OTM as a fallback. Highways were identified by road class and received their own emission factor independent of their speed limit.~~

~~The resulting dataset is referred to in Figure 1 as *bottom-up* emissions. We then scaled bottom-up emissions with national totalsand gridded with 0.05° x 0.1° resolution to generate our final results for CAMS-REG v8.1. Additionally, we applied higher resolution urban grids for a comparison with city inventories.~~

2.4 Comparison to CAMS-REG v4.2 method

Based on the introduced methodology and datasets, CAMS-REG v8.1 includes the first major update on the spatial distribution of road emissions since v4.2, which was published in 2022 (Kuenen et al. (2022)). At the core, both methods are based on the activity data distribution from OTM in combination with OSM shapes and a methodology to fill up missing data (see Figure 1. In v4.2, the link-based activity data is first filled up by taking an average of the road type within each corresponding NUTS3 area, or, if not available, a global average. Then, emission factors ~~are~~ from COPERT were applied, and the spatial distribution ~~is~~ was gridded to a fixed grid of 0.05° x 0.1° degrees resolution. Lastly, empty grid cells with no available data; for example in non-EU Eastern European countries were filled up using a regression approach based on Landscan population density data, per road class. The resulting spatial dataset was then used to distribute national total emissions per country.

The spatial distribution, as well as large parts of the data handling within this version, is dependent on the exact grid definition, and an increase of grid resolution is not possible beyond the resolution of Landscan for large parts of the study area. If city scale or higher resolution inventories need to be derived from this approach, additional tools and proxies are always necessary. So far, this has been done by either tools such as UrbEm (Ramacher et al. (2021)), or proxies such as local road geometries or other distributions.

2.5 Evaluation methods

Here, we describe the evaluation of our intermediate results and the final dataset.

~~Gap-filling~~ RF gap filling model performance was tested with 10% reserved test data and the coefficient of determination (R^2), mean absolute error (MAE) and root mean square error ($RMSE$) as performance metrics.

Due to the scaling with national emission totals (Figure 1), national totals of the final emission datasets are bound to be similar to previous versions on a national level, despite the base year difference of 2017 (v4.2) and 2018 (v8.1). National totals used to scale emissions for the two CAMS-REG versions used may be slightly different due to updates in the national methodologies.

We then compare our results to the 8 made-available city scale inventories (see Chapter 2.1.4). For this, the boundary shapes of the city inventories were used to extract annual emission totals from CAMS-REG versions, which were then compared to the city inventories. The areas of these boundary shapes included regions of different sizes, which are available as described in Section 5. The largest city boundary was for the Paris-Ile de France region (ca. 12470 km^2).

For a wider picture, we included 16 other European cities for a comparison between CAMS-REG versions. As different national totals between the CAMS-REG versions can also affect city allocation, comparisons were also done on a *'Share of urban emissions to national totals'* basis.

For London and Milan, in addition to city totals also a high-resolution spatial distribution is available from LAEI, which enables a case-study comparison of the spatial distribution between different methods. We gridded our new distribution to the LAEI grid, and performed an UrbEm downscaling of CAMS-REG v4.2 to the same grid. We then performed a spatial correlation between the LAEI inventory and both CAMS inventories, as well as spatial cross-correlation. Spatial cross-correlation has been proposed by Chen (2015) as a way to assess the relationship between one variable at a location and another variable at other locations. It is an extension of the commonly used Moran's I index for spatial autocorrelation over another variable (Chen (2015)). Possible values reach from -1 to 1, with -1 denoting perfect negative cross-correlation, 1 denoting perfect positive cross-correlation, and 0 denoting no cross-correlation.

3 Results

In this section, we give the performance of the trained AADT models~~, as well as~~. The gap filling of AADT data is done with the aim of generating a complete activity dataset for the later emission calculation. The unit of AADT is vehicles per day. In the following, we give an overview on the resulting emissions dataset. We then compare the total results with emissions from local inventories. We further included a comparison of the spatial distribution of local inventories for London and Milan inventories against CAMS-REG.

3.1 Model performance

~~Gap-filling~~ RF Gap filling performance was dependent on model and road type. Model 2 (see Table 3) reaches $R^2 = 0.63$, whereas Model 1 (trained also with OTM data) reaches $R^2 = 0.74$. This performance decreases along decreasing road size, with primary roads achieving $R^2 = 0.64$, secondary roads $R^2 = 0.51$ and tertiary roads $R^2 = 0.44$ for Model 2 and $R^2 = 0.47$, $R^2 = 0.39$ and $R^2 = 0.29$ for Model 1, respectively. The MAE and RMSE for Model 2 are $MAE = 2028.3$ and $RMSE = 4407.2$. Errors increase with road class (highway $MAE = 5182.3$, tertiary roads $MAE = 619.7$). For the Model 1, errors were $MAE = 1570.2$ and $RMSE = 3702.4$.

For motorways, the highway models achieve $R^2 = 0.66$ and $R^2 = 0.76$ for Highway Model 2 and Highway Model 1, respectively.

We further calculated variable importance according to ~~Pulugurtha and Mathew (2021)~~Ishwaran and Lu (2019). For Model 2, the most important predictor variables were the OSM highway key, and then population density and land use class. For Model 1, the OTM road class was the most important predictor, followed by population density and the OSM highway key. For both models, OSM data on maximum road speed were of low importance.

Figure 2. Performance of ~~AADT–gap–filling~~ RF AADT–gap filling model ~~2~~ 1 (for use in EU countries), by road class ~~.~~for the year 2018.

3.2 Validation of activity dataset with measured dataset

We compared our complete activity dataset to a recent dataset published by Bonnemaizon et al. (2025). This dataset provides measured, annual averaged traffic data for a number of European cities between 2015 to 2024 depending on availability. The locations of traffic counts were matched with the OSM network by Hidden Markov Models (in case of line geometries) and nearest neighbour approach (in case of point data). The performed matching with the OSM network provided a straightforward way for comparison with our dataset.

For the year 2018, data was available for 13 cities, for a total number of 6219 traffic counting stations. The number of stations per city varied strongly, with the smallest number in Copenhagen (n = 17) and the largest number in London (n = 2207). Motorways (motorway + trunk) accounted for 1056 stations, primary (primary + primary_link) for 1570 stations, secondary (secondary + secondary_link) for 825 stations, tertiary (tertiary + tertiary_link) for 1217 stations and residential (residential + service + living_street) for 1430 stations. Descriptive statistics were (min: 5; 25% quantile: 3775; median: 10288; mean: 15405; 75% quantile: 18530; max: 163045), compared to our generated dataset: (min: 0; 25% quantile: 221; median: 1449; mean: 4508; 75% quantile: 3841; max: 107656). Frequent high values for residential streets in the dataset from Bonnemaizon et al. (2025) hint to a possible mismatch of some sensor locations. For example, the mean AADT of service

roads was 21572, which is likely be explained by a location mismatch. Bonnemaizon et al. (2025) are aware of this source of error.

Correlation between both datasets was found at $R^2 = 0.17, p < 0.01$ with a slope of $b = 0.84$. Bonnemaizon et al. (2025) mention higher certainty of their matching algorithm with reduced distance between sensor and matched roads. We therefore performed the analysis again in two steps. First, we only took into account sensors with matched roads close to the sensor location ($distance < 1m; n = 3476$), as suggested by the authors to help reduce matching uncertainty. Secondly, we only took into account sensors that came with a matched line geometry, increasing location certainty even further ($n = 2249$). In the first case, correlation increased to $R^2 = 0.20$, and in the latter case to $R^2 = 0.38$, which is still below $R^2 = 0.5$, which is seen satisfactory by Bonnemaizon et al. (2025). Our RF gap filling methodology did not seem to negatively impact correlation with the validation dataset. Using only gap filled data, correlation increases from $R^2 = 0.17$ to $R^2 = 0.27$.

The statistics show a large underestimation in our dataset. The underestimation is the largest for small roads (by factors of 210 for service roads and 18 for residential roads, with likely reasons as outlined before). Amongst other road classes, there is an underestimation of our dataset by a factor ranging from 2-6 for different per road class. As road traffic volume in Europe still generally continues to grow (EEA (2024)), a part of the underestimation may be explained by OTM's 2015 base year compared to the measured data from 2018. However, there seems to exist an additional systematic underestimation of AADT values in the OTM dataset for urban areas compared to the measured data.

Our findings of a substantial increase in correlation with decreasing geolocation uncertainty is in line with previous results by Bonnemaizon et al. (2025), who explain weak correlation with a validation datasets due to uncertainty in geolocation of sensor data. For the final purpose of generating a relative spatial distribution, the low correlation values between between our dataset and the measured data are a limitation to our results, highlighting the need for an updated and consistent AADT dataset in Europe.

3.3 Resulting emissions and datasets

As absolute emissions per country were scaled by reported country totals, emissions on the country levels between CAMS-REG versions differ only due to changes in country reporting methodology for a given year. Figure 3 gives an overview of changes in country totals. The NOx domain total for CAMS-REG v4.2 was 3.67×10^9 kg, and 4.03×10^9 kg in v8.1 ~~. The largest~~ for 2018. The largest absolute increases in reported emissions were for Turkey and Russia and the largest decrease, for Poland. We further calculated the changes of urban core emissions between the versions. Urban core areas were taken from the Global Human Settlement Layer (Pesaresi et al. (2024)). Throughout the domain, urban emissions increased by 26%. Large increases can be found in non-EU countries such as Georgia, Russia, Serbia and Turkey (4). Strong decreases can also be found, for example in Estonia and Finland, which occur due to the spatial redistribution of emissions.

A distribution of absolute emissions by OSM class is given in Table 5. Over 65% of all emissions occurs on trunk, motorway and primary roads, with another roughly 15% of emissions occurring on secondary and tertiary roads. Despite their large number, smaller roads only account to 3.7% of all emissions in the domain.

OSM class	Absolute emissions (NOx kg/y)	% of domain total
trunk	1352338521	19.83
motorway	1552731489	22.77
motorwaylink	82878332	1.22
primarylink	16905031	0.25
primary	1478008871	21.68
secondary	1056816802	15.50
tertiary	1020387335	14.96
tertiarylink	3872399	0.05
residential	120131847	1.76
livingstreet	4254268	0.06
track	80517356	1.18
service	43686624	0.64

Table 5. Distribution of absolute and relative shares of NOx emissions in the domain by OSM class

Figure 5 gives an overview of the spatial distribution for France, Russia and Ukraine. France shows a higher allocation to urban centres, with a largely unchanged spatial distribution for the rest of the country. Russia and Ukraine show vastly different spatial distributions, with more emissions attributed to cities, visible urban centres and highways, as well as a reduction in artifacts.

Figure 3. ~~Comparison~~ Percentage difference of ~~country totals~~ NOx emission estimations between CAMS-REG versions ~~4.2 and~~ 8.1 and 4.2 for the year 2018. Positive values denote higher estimations in version 8.1. Labels are absolute emissions (t/y).

Figure 4. Percentage difference of urban NOx emission estimations between CAMS-REG versions 8.1 and 4.2 for the year 2018. Positive values denote higher estimations in version 8.1. Labels are absolute emissions (t/y). *We cut the value for Moldova (MDA) to preserve scale (409%)

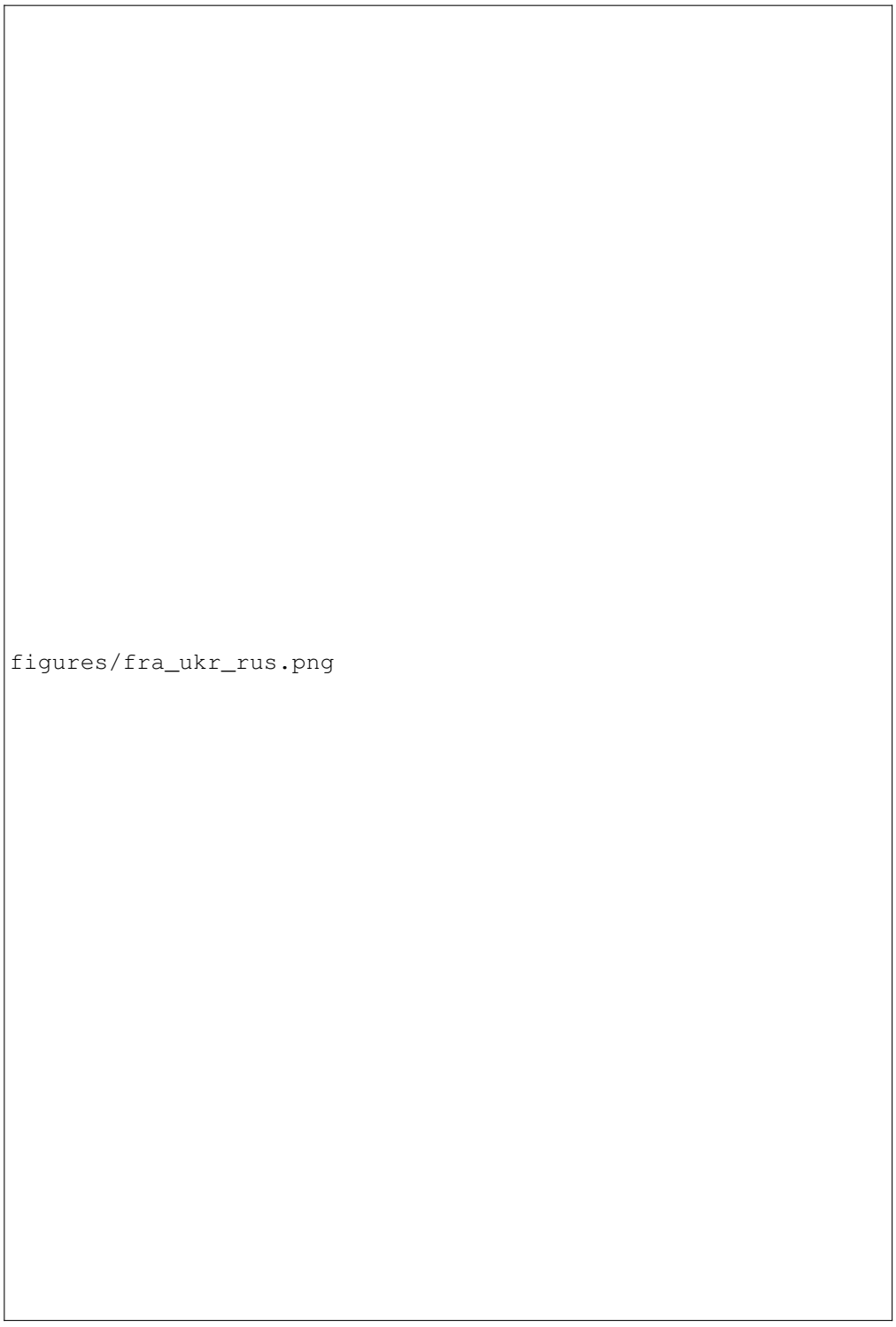

Figure 5. Spatial distribution of road emissions for France (a), Ukraine (b) and western Russia (c) between CAMS-REG versions 4.2 and 8.1. Please note the different colour scales on the left and right side plots of (b).

3.4 Comparison with city inventories

As city inventories compiled by local institutes or authorities are prepared with the most detailed local knowledge, they are here seen as the gold standard with which to compare our European emission dataset. Nonetheless, there could be also significant errors in the spatial distribution and absolute emission estimates of local city inventories. We compared the absolute NOx estimations for CAMS-REG v4.2 and CAMS-REG v8.1 to the data of eight cities for which independent city inventories was made available to us (Chapter 2.1.4).

Absolute emission comparisons are given in Figure 6. The mean absolute percentage error of CAMS-REG v4.2 in comparison with local city inventories was 35%, with the smallest error for Paris (18%) and the largest error for Bologna (53%). All cities showed underestimations in CAMS-REG v4.2 compared with local inventories, which we explain by the generally too-low urban AADT values from OTM-based estimation (Section 3.2), as well as the more basic gap filling approach employed in v4.2.

Using CAMS-REG v8.1 emissions, the average difference was ~~19~~18% compared with the city scale inventories. All estimates improved except for Helsinki, and improvements were largest for Barcelona and Milan. CAMS-REG v8.1 ~~also~~ largely underestimates emissions for ~~most cities, especially Helsinki (56%)and Bologna (40%). Milan was the only city with overestimated emissions (by 7~~Helsink (-26%), and overestimates emissions for Milan (+26%).

Higher emission totals in urban areas were also estimated for most of the cities for which no locally-made city inventory was available for comparison (see right side of Figure 6). Here, ~~Istanbul, Saint Petersburg and Moscow~~ Kiev and Istanbul saw strong increases compared to CAMS-REG v4.2 ~~. These cities now have comparable emissions to their Western European counterparts. Past large underestimations of their emissions could be attributed to the inability of the previous method to handle missing data for Russia and other European countries~~(+84%, +360%, respectively).

Absolute emission estimates are impacted not only by the spatial distribution of country totals, but also by the country totals themselves. Between CAMS-REG v4.2 and v8.1, there has been a change in reported country totals for some countries (see Section 3.3). To understand the separate effects of the spatial distribution, we compared the relative shares of urban NOx emission to country totals between both versions (Figure 7). Here, the relative shares of urban emissions increased for most cities, with the exception of Helsinki, Athens and Porto. The case of Helsinki is the only city showing a decreasing performance compared to the local inventory. In CAMS-REG v4.2, a large proportion of Finnish national emissions were allocated to the capital city (Figure 7), in line with the population-based method used (Section 2.4). In Finland, over 30% of the population lives in Helsinki, which is a high value for Europe (Eurostat (2024)).

The absolute emission changes in each city follow the changes of relative urban emission shares ~~well~~ (Figures 6, 7). Therefore, the changes between CAMS-REG versions can be attributed to the updated spatial distribution. Across all cities, the ~~median~~ mean increase of urban emission share was ~~24%. Moscow and St. Petersburg saw very~~ 34%. Istanbul, Milan and Kiev saw high increases of relative shares (~~453% and 288~~160% and 100%, 90%, respectively) ~~, due to the large changes in the spatial distribution Russia~~ (see Figure 5).

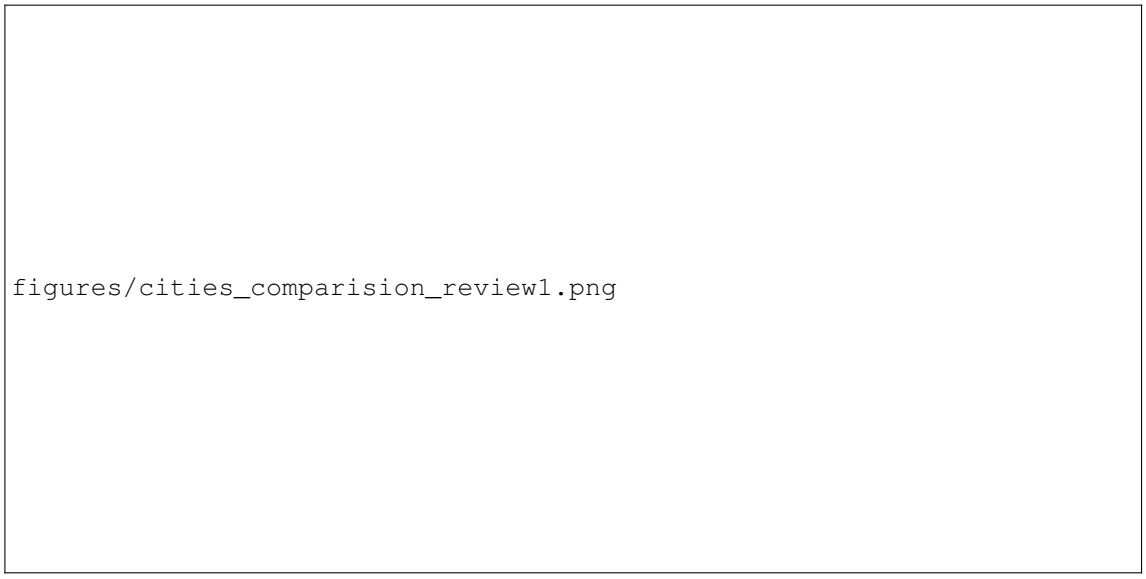

Figure 6. Comparison of annual NOx emissions from road transport of CAMS-REG 4.2, CAMS-REG 8.1 and local city inventories for ~~2018 (left side).~~ 2018. Years covered by local inventories are given in Table 2. ~~A series of other European cities for which no local inventory was available are given as reference (right side).~~ Note that the areas covered by city boundary shapes vary substantially. The city boundaries used for this comparison are provided in Section 5.

Figure 7. Share of NOx urban emissions to national totals between CAMS-REG versions 4.2 and 8.1 for road transport in 2018.

3.5 Spatial distribution London

Figure 8 compares the spatial distribution for London between the urban inventory (LAEI) and the previous and recent CAMS-REG estimates in 1x1km resolution, normalized by total urban emissions. The original 0.05° x 0.1° degrees resolution of the CAMS-REG v4.2 dataset was downscaled to 1x1km using the UrbEm tool (Ramacher et al. (2021)). In the local and CAMS-REG v8.1 estimates, major roads such as the outer M25 ring or the North Circular Ring are clearly visible. Compared to the local inventory, CAMS-REG v8.1 seems to overestimate emissions at certain highways. Spatial correlation between the local inventory and CAMS-REG v8.1 reaches $R^2 = 0.62$, with $p < 0.001$, and spatial cross-correlation was calculated at $I = 0.06, p < 0.001$. Spatial correlation between the local inventory and CAMS-REG v4.2 reaches $R^2 = 0.33$, with $p < 0.001$ (downscaled with UrbEm), and spatial cross-correlation at $I = 0.11, p < 0.01$. Spatial autocorrelation of the local inventory (Moran's I) is $I = 0.07$.

Figure 8. Comparison of the spatial distribution of urban NOx road transport emissions for London with local (LAEI, base year 2019) and CAMS-REG inventories (1 km resolution, base year 2018).

3.6 Spatial distribution Milan

To derive a local inventory for Milan, Marongiu et al. (2024) started with a complete road graph that had been simplified to include the fewest oriented arcs capable of representing the most significant motions. Data from motorway companies, ANAS (National Autonomous Roads Corporation) and local authorities of vehicular traffic data (from traffic counters, cameras, passages at toll booths and motorway barriers) were used in a flow assignment model for estimating the origin and destination matrices over the entire network. To extend the spatial extension of the detailed estimates and avoid the calculation of diffuse traffic emissions, the simplified road graph was used to train and test (with a data split of 75% and 25%) a RF model for predicting emissions by road arch, according to the following variables: vehicles (cars, high-duty, low-duty and motorcycles) and 12 variables describing road characteristics. The ML performances are confirmed for training and testing in the range of an R^2 of: 0.8 - 0.6 for high duty, 0.9 - 0.8 for light duty, 0.8 – 0.7 for motorcycles and 0.9 - 0.7 for passenger cars. For a detailed description of the Milan local inventory, see Marongiu et al. (2024).

An overview of the spatial distribution of road transport emissions from the local Milan inventory and CAMS-REG versions are shown in Figure 9. Like to London, the CAMS-REG v4.2 inventory was downscaled using the UrbEm tool to match the spatial resolution of the local inventory. The Milan local inventory places most emissions on the outer ring road of Milan, while also estimating emissions for small roads. CAMS-REG v8.1 places more emissions on the inner ring road, and less emissions on local roads in general. Ring roads and highways are less visible in CAMS-REG v4.2. Spatial correlation between the local inventory and CAMS-REG v4.2 (downscaled with UrbEm) inventories reaches $R^2 = 0.25$, with $p < 0.001$, and spatial cross-correlation is at $I = 0.08$ with $p < 0.001$. The spatial correlation between the local inventory and CAMS-REG v8.1 is $R^2 = 0.59$, with $p < 0.001$, and spatial cross-correlation reaches $I = 0.06$ with $p < 0.001$. Moran's I (spatial autocorrelation) for the local inventory is at $I = 0.07$. Low spatial autocorrelation is captured in both versions of CAMS-REG.

Figure 9. Comparison of the spatial distribution of urban NOx road transport emissions for Milan with local (base year 2019) and CAMS-REG inventories (0.01° x 0.007° resolution~~).~~, base year 2018.

3.7 Comparison between speed-based and land-use–based allocation of urban and rural roads

Our emission factor set differentiates between urban roads, rural roads and highways. The allocation to highways was straightforward, and is done via OSM classes (see Table 3).

We further tested the impact of the methodology to determine whether a road receives an urban or rural emission factor. In the previous version of this dataset, the emission factor of roads (or of gridded road proxies) was determined by land use class or population density (Kuenen et al. (2022)). In this update, we set the road status by the maximum driving road speed allowed on the road (see Chapter 1). This was thought to better represent real-world emission factors than the land use class around the road and align the method with the approach of the COPERT emission inventory model (Ntziachristos et al. (2009)). For the first method, we gave an urban status to roads within an urban CORINE grid cell (CORINE categories 1.1.1 or 1.1.2). For the second method, we attributed urban status for roads with a maximum speed of $\leq$ 50km/h.

Figure 10 shows a comparison of urban and rural allocation for both methods. It is visible that the allocation by maximum speed leads to more emissions being classified as urban, both inside and outside urban areas. These roads subsequently receive urban (and therefore higher) emission factors compared to rural roads. Over the whole domain, the ratio of urban to rural emissions is greatly different for both approaches. Using a speed-based approach, the ratio between urban and rural emissions is approximately 4:1, whereas it is approximately 1:2 when using a land-use–based approach. This leads to a wider use of higher, urban emission factors in the speed-based approach (see Figure 10). We tested the impact of both methods on overall city emissions (Section 3.4). For each city, the difference was very small and typically less than 1% of the final result.

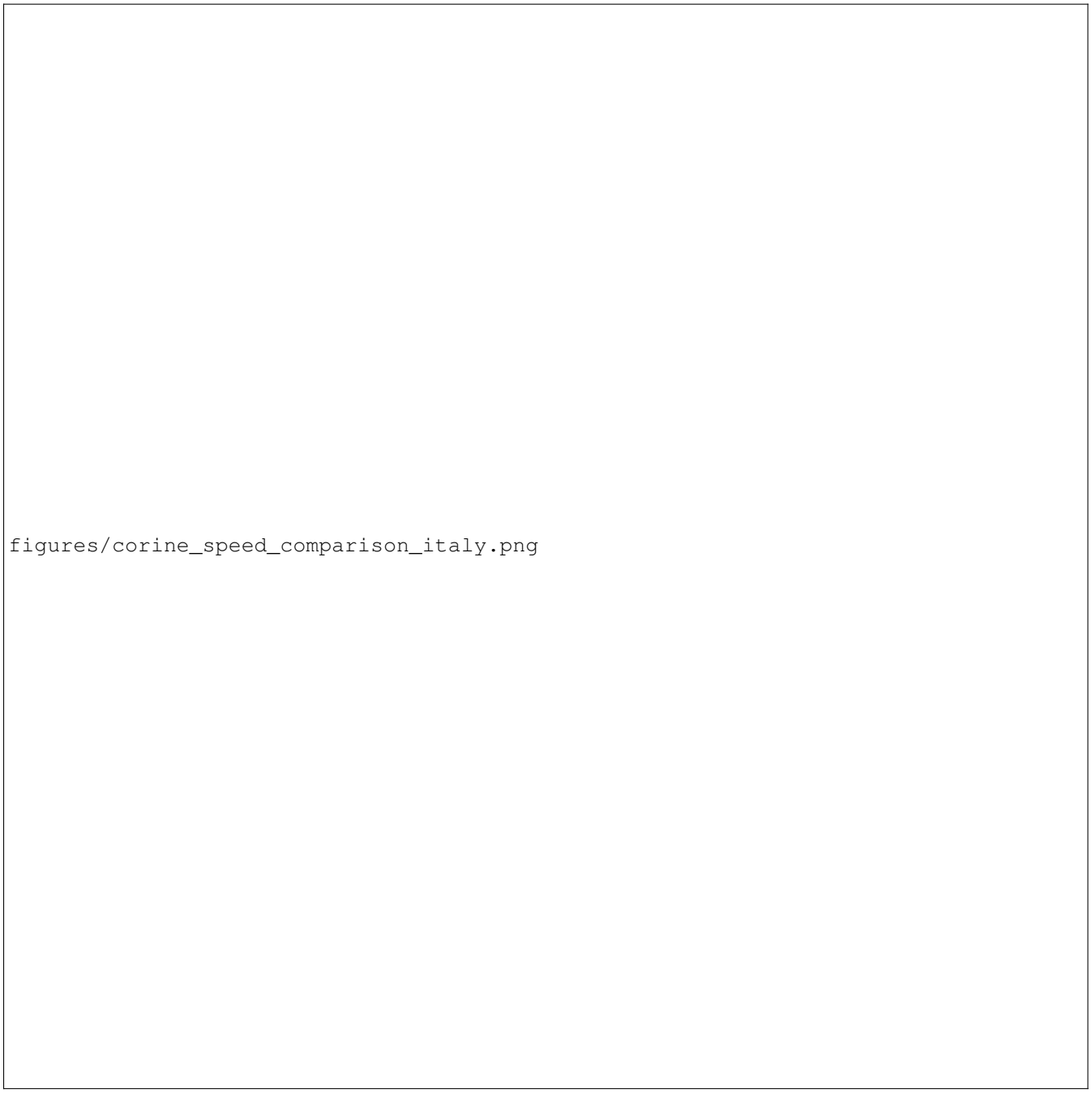

Figure 10. Comparison between speed-based and land-use–based urban/rural distinction for Italy ~~.~~for 2018.

This study presents a new spatial distribution for road transport emissions in Europe to improve CAMS-REG emission inventories. The motivation for a new spatial distribution was three-fold. Firstly, comparisons with urban inventories showed frequent underestimations of urban NOx emissions by CAMS-REG v4.2. Secondly, work-intensive methods such as the use of downscaling tools or urban proxies had to be employed to create urban inventories with high resolutions needed for detailed exposure studies or urban air quality management (Venter et al. (2023); Gulia et al. (2015); Yang et al. (2019)) from existing emission inventories. Thirdly, spatial distributions for countries in Eastern Europe were faulty and suffered from artifacts due to the lack of available data. To overcome these challenges, a vector-based methodology was applied, in which emissions are calculated for all roads on a link level first, and then flexibly gridded to the required resolution. A major part of the work was the gap filling of missing activity data to create a coherent dataset. This was done using random forest models and OTM, OSM with land use and population density data, and a spatial approach for the smallest roads. Comparing the resulting dataset to independent city inventories revealed an improved NOx emission allocation in ~~6 out of 7~~ seven out of eight cities for which data was available, and an allocation of emissions to non-EU Eastern European cities comparable to their EU counterparts. Case study comparisons of spatial distributions in London and Milan lead to similar results. Compared at ~~a~~ an approximately 1-km scale, in both cases spatial correlation increased from around $R^2 \approx 0.3$ using CAMS-REG v4.2 and UrbEm downscaling to $R^2 \approx 0.6$ with CAMS-REG v8.1. As the random forest models rely on OSM data for training (for example for road class information) as well as for road geometries, the community character of OSM is introducing an uncertainty source. Even though OSM data has been found reliable in previous studies (Demetriou (2016), Moradi et al. (2021)), we found the data availability of attributes low, and different units (mp/h and km/h) in maximum speed data needed to be translated, ~~giving space to further errors~~with possible error room for different labelling conventions.

We trace the previously too-low allocation of emissions to urban centres back to the incomplete and more simple gap filling performed in CAMS-REG v4.2. Here, a larger share of emissions was attributed to highways, as their data availability was higher. The complete estimation of activity data, which now also includes small roads, led to the increase of urban emission allocation and a better comparison with urban inventories. Chapter 3.7 further shows a largely different ratio of urban and rural emissions based on the allocation method. As all emissions were scaled by national totals, this did not change overall emissions, but also contributed to a higher share of emissions within city centres. When using land-use–based methods, we found a proportion of urban centres outside urban categories (Corine code 1.1.1 or 1.1.2), leading to a likely misallocation of roads. Further, higher emission factors were strictly limited to urban centres, even though many roads in rural areas may also be operated on with low speeds or high congestion levels, which will lead to higher emissions for these roads ~~.~~ (see Kelly (2025); Li et al. (2023) for related research).

The selection of emission factors for road links based on speed limit or congestion status is seen as an important step towards an improved spatial distribution.

For consistency with previous CAMS-REG versions, we based our approach on the OTM dataset, which has also been used in CAMS-REG v4.2 as the main source of activity data. The OTM dataset was published in 2016, after the use in a number of

pilot regions as part of the EU-funded Open Transport Net project (Jedlicka et al. (2016)). The results could be improved by the addition of more and more up-to-date datasets. These could be integrated based on their OSM-ID in a hierarchical fashion, in which measured and newer data takes precedence over modelled or older data, while taking care to produce a harmonic dataset. This integration work will be a major task for the future.

The published dataset is to our knowledge the first attempt to provide coherent vector-based road emissions for most roads in Europe, which can be gridded in any resolution to serve as a base for detailed urban air quality analysis and exposure studies. The dataset can be used in cities to provide a starting point for their own work on an inventory and as a reference point for comparison, which will help to improve our estimates as well. Future work is planned to expand the spatial distribution to ultrafine particles (UFP). A related recent work was done by Shen et al. (2024). Here, the authors used RF to predict Europe-wide AADT values on a 5 metre grid resolution based on measurement stations in six European countries.

Whereas this update shows strong improvements in the allocation of more emissions to urban centres, average urban emissions are still underestimated by around 20% compared to locally compiled city inventories (see Chapter 3.4). Moreover, locally compiled inventories themselves may underestimate traffic emissions, as shown in Karl et al. (2024) through a comparison with traffic sensors for PM2.5 in Hamburg. The current methodology only differentiates between *'urban'* and *'rural'* emissions, without taking into account additional information on flow and congestion, despite their important effects on emissions rates (Xue et al. (2013); Treiber et al. (2007); Barth and Boriboonsomsin (2008)). We therefore see the combination with traffic flow and congestion information as an important next step for future work. Lastly, the performance of air quality simulations in reproducing measured urban concentrations based on these changes will have to be assessed as part of a future perspective.

5

Data is available in vector format and gridded raster format at a resolution of 0.05° x 0.1°. The vector dataset includes all information necessary to derive a gridded spatial distribution for different resolutions and extents. Fields include traffic volume, road class, road length, country code, urban/rural/highway category, and total emissions per vehicle type, as well as total vehicle kilometre per vehicle type. The raster dataset is gridded per vehicle type and urban/rural/highway category. Moreover, the city boundary shapes used to compare local and CAMS-REG inventories are included. Detailed information on each field of the vector file are given in an included Readme file.

Data can be found at https://doi.org/10.5281/zenodo.15688723 (Hohenberger et al. (2025)).

. All authors contributed significantly to this research. TH, JK and AV conceptualized the study. TH wrote the model code and generated the results. MM compliled the emissions factors. TH, JK, AV and MM contributed to the analysis. MG, MR, AM, GL, GF and AK compiled the city inventories for comparison and reviewed the manuscript.

. The contact author has declared that none of the authors has any competing interests.

. We are grateful to Ingrid Super for providing comments on the initial manuscript. This work was supported by the European Union's Horizon Europe (2021–2027) research and innovation programme under project EASVOLEE, grant agreement no. 101095457.

References

Achebak, H., Garatachea, R., Pay, M. T., Jorba, O., Guevara, M., Pérez García-Pando, C., and Ballester, J.: Geographic sources of ozone air pollution and mortality burden in Europe, Nature Medicine, 30, 1732–1738, https://doi.org/10.1038/s41591-024-02976-x, 2024.

Airparif: Paris Emissions 2019, https://www.airparif.fr/en/surveiller-la-pollution/les-emissions, 2025.

Antonczak, B., Fay, M., Chawla, A., and Rowangould, G.: Estimated Roadway Segment Traffic Data by Vehicle Class for the United States: A Machine Learning Approach, arXiv preprint arXiv:2502.05161, 2025.

Apronti, D., Ksaibati, K., Gerow, K., and Hepner, J. J.: Estimating traffic volume on Wyoming low volume roads using linear and logistic regression methods, Journal of Traffic and Transportation Engineering (English Edition), 3, 493–506, https://doi.org/10.1016/j.jtte.2016.02.004, 2016.

Arpa Emilia-Romagna: Inventario regionale emissioni in atmosfera (INEMAR) - Dati Arpae, https://dati.arpae.it/dataset/inventario-emissioni-aria-inemar, 2025.

Baffoe-Twum, E., Asa, E., and Awuku, B.: Estimation of annual average daily traffic (AADT) data for low-volume roads: a systematic literature review and meta-analysis, Emerald Open Research, 1, https://doi.org/10.1108/EOR-05-2023-0010, 2023.

Barnes, J. H., Chatterton, T. J., and Longhurst, J. W. S.: Emissions vs exposure: Increasing injustice from road traffic-related air pollution in the United Kingdom, Transportation Research Part D: Transport and Environment, 73, 56–66, https://doi.org/10.1016/j.trd.2019.05.012, 2019.

Barth, M. and Boriboonsomsin, K.: Real-World CO2 Impacts of Traffic Congestion, Transportation Research Record, 2008.

Bhaduri, B., Bright, E., Coleman, P., and Dobson, J.: LandScan, Geoinformatics, 5, 34–37, 2002.

Bonnemaizon, X., Ciais, P., Zhou, C., Shi, Q., Mittakola, R. T., Goldmann, C., Ben Arous, S., Megel, N., and Davis, S. J.: Harmonized Annual Averaged Traffic Data at Street Segment Level for European Cities, Scientific Data, 12, 1365, https://doi.org/10.1038/s41597-025-05698-y, 2025.

Chang, Y., van Strien, M. J., Zohner, C. M., Ghazoul, J., and Kleinschroth, F.: Effects of climate, socioeconomic development, and greening governance on enhanced greenness under urban densification, Resources, Conservation and Recycling, 206, 107 624, https://doi.org/10.1016/j.resconrec.2024.107624, 2024.

Chen, Y.: A New Methodology of Spatial Cross-Correlation Analysis, PLOS ONE, 10, e0126 158, https://doi.org/10.1371/journal.pone.0126158, 2015.

Colette, A., Collin, G., Besson, F., Blot, E., Guidard, V., Meleux, F., Royer, A., Petiot, V., Miller, C., Fermond, O., Jeant, A., Adani, M., Arteta, J., Benedictow, A., Bergström, R., Bowdalo, D., Brandt, J., Briganti, G., Carvalho, A. C., Christensen, J. H., Couvidat, F., D'Elia, I., D'Isidoro, M., Denier Van Der Gon, H., Descombes, G., Di Tomaso, E., Douros, J., Escribano, J., Eskes, H., Fagerli, H., Fatahi, Y., Flemming, J., Friese, E., Frohn, L., Gauss, M., Geels, C., Guarnieri, G., Guevara, M., Guion, A., Guth, J., Hänninen, R., Hansen, K., Im, U., Janssen, R., Jeoffrion, M., Joly, M., Jones, L., Jorba, O., Kadantsev, E., Kahnert, M., Kaminski, J. W., Kouznetsov, R., Kranenburg, R., Kuenen, J., Lange, A. C., Langner, J., Lannuque, V., Macchia, F., Manders, A., Mircea, M., Nyiri, A., Olid, M., Pérez García-Pando, C., Palamarchuk, Y., Piersanti, A., Raux, B., Razinger, M., Robertson, L., Segers, A., Schaap, M., Siljamo, P., Simpson, D., Sofiev, M., Stangel, A., Struzewska, J., Tena, C., Timmermans, R., Tsikerdekis, T., Tsyro, S., Tyuryakov, S., Ung, A., Uppstu, A., Valdebenito, A., Van Velthoven, P., Vitali, L., Ye, Z., Peuch, V.-H., and Rouïl, L.: Copernicus Atmosphere Monitoring Service – Regional Air Quality Production System v1.0, https://doi.org/10.5194/egusphere-2024-3744, 2024.

Copernicus: Corine Land Cover, https://land.copernicus.eu/en/products/corine-land-cover, 2018.

Cornell Local Roads Program: Basics of a Good Road, 2014.

Das, S. and Tsapakis, I.: Interpretable machine learning approach in estimating traffic volume on low-volume roadways, International Journal of Transportation Science and Technology, 9, 76–88, https://doi.org/10.1016/j.ijtst.2019.09.004, 2020.

Demetriou, D.: Uncertainty of OpenStreetMap data for the road network in Cyprus, p. 968806, https://doi.org/10.1117/12.2239612, 2016.

Denier van der Gon, H., Gauss, M., Granier, C., Arellano, S., Benedictow, A., Darras, S., Dellaert, S., Guevara, M., Jalkanen, J.-P., Krueger, K., Kuenen, J., Liaskoni, M., Liousse, C., Markova, J., Prieto Perez, A., Quack, B., Simpson, D., Sindelarova, K., and Soulie, A.: Documentation of CAMS emission inventory products, https://doi.org/10.24380/Q2SI-TI6I, publisher: Copernicus Atmosphere Monitoring Service, 2023.

EEA: Sustainability of Europe's mobility systems 2024, https://www.eea.europa.eu/en/analysis/publications/sustainability-of-europes-mobility-systems, 2024.

El-Harbawi, M.: Air quality modelling, simulation, and computational methods: a review, Environmental Reviews, 21, 149–179, 2013.

EMEP Centre of Inventories and Predictions: Air pollutant emissions data viewer, https://www.eea.europa.eu/en/topics/in-depth/air-pollution/air-pollutant-emissions-data-viewer-1990-2023, 2023.

Eurostat: Urban-rural Europe - demographic developments in cities, https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Urban-rural_Europe_-_demographic_developments_in_cities, 2024.

Finnish Meteorological Institute: Homepage, https://en.ilmatieteenlaitos.fi/, 2025.

Franco, V., Kousoulidou, M., Muntean, M., Ntziachristos, L., Hausberger, S., and Dilara, P.: Road vehicle emission factors development: A review, Atmospheric Environment, 70, 84–97, https://doi.org/10.1016/j.atmosenv.2013.01.006, 2013.

Fu, X., Xiang, S., Liu, Y., Liu, J., Yu, J., Mauzerall, D. L., and Tao, S.: High-resolution simulation of local traffic-related NOx dispersion and distribution in a complex urban terrain, Environmental Pollution, 263, 114 390, https://doi.org/https://doi.org/10.1016/j.envpol.2020.114390, 2020.

Geilenkirchen, G., Bolech, M., Hulskotte, J., Dellaert, S., Ligterink, N., and van Eijk, E.: Methods for calculating the emissions of transport in the Netherlands, Tech. rep., Rijksinstituut voor Volksgezondheid en Milieu RIVM, https://doi.org/10.21945/RIVM-2024-0023, 2024.

Grange, S. K., Farren, N. J., Vaughan, A. R., Rose, R. A., and Carslaw, D. C.: Strong temperature dependence for light-duty diesel vehicle NO x emissions, Environmental Science & Technology, 53, 6587–6596, 2019.

Greater London Authority: London Atmospheric Emissions Inventory (LAEI) 2019, https://data.london.gov.uk/dataset/london-atmospheric-emissions-inventory--laei--2019, 2023.

Gu, Y., Wong, T. W., Law, C., Dong, G. H., Ho, K. F., Yang, Y., and Yim, S. H. L.: Impacts of sectoral emissions in China and the implications: air quality, public health, crop production, and economic costs, Environmental Research Letters, 13, 084 008, 2018.

Guevara, M., Tena, C., Porquet, M., Jorba, O., and Pérez García-Pando, C.: HERMESv3, a stand-alone multi-scale atmospheric emission modelling framework – Part 2: The bottom–up module, 13, 873–903, https://doi.org/10.5194/gmd-13-873-2020, 2020.

Gulia, S., Shiva Nagendra, S., Khare, M., and Khanna, I.: Urban air quality management-A review, Atmospheric Pollution Research, 6, 286–304, https://doi.org/10.5094/APR.2015.033, 2015.

Gurram, S., Stuart, A. L., and Pinjari, A. R.: Agent-based modeling to estimate exposures to urban air pollution from transportation: Exposure disparities and impacts of high-resolution data, Computers, Environment and Urban Systems, 75, 22–34, https://doi.org/10.1016/j.compenvurbsys.2019.01.002, 2019.

Han, D. C.: Prediction of Traffic Volume Based on Deep Learning Model for AADT Correction, Applied Sciences, 14, 9436, https://doi.org/10.3390/app14209436, 2024.

Hohenberger, T. L., el Malki, M., Ramacher, M. O. P., Guevara, M., and Kuenen, J.: Spatial distribution road transport emissions for CAMS-REG v8.1, https://doi.org/10.5281/zenodo.15688723, 2025.

Hopke, P. K., Feng, Y., and Dai, Q.: Source apportionment of particle number concentrations: A global review, Science of The Total Environment, 819, 153 104, https://doi.org/https://doi.org/10.1016/j.scitotenv.2022.153104, 2022.

Int Panis, L., Broekx, S., and Liu, R.: Modelling instantaneous traffic emission and the influence of traffic speed limits, Science of The Total Environment, 371, 270–285, https://doi.org/10.1016/j.scitotenv.2006.08.017, 2006.

Isakov, V., Touma, J. S., Burke, J., Lobdell, D. T., Palma, T., Rosenbaum, A., and kÖzkaynak, H.: Combining regional-and local-scale air quality models with exposure models for use in environmental health studies, Journal of the Air & Waste Management Association, 59, 461–472, 2009.

Ishwaran, H. and Kogalur, U.: Fast Unified Random Forests for Survival, Regression, and Classification (RF-SRC), https://cran.r-project.org/package=randomForestSRC, r package version 3.3.1, 2024.

Ishwaran, H. and Lu, M.: Standard errors and confidence intervals for variable importance in random forest regression, classification, and survival, Statistics in Medicine, 38, 558–582, https://doi.org/10.1002/sim.7803, 2019.

Jedlicka, K., Hajek, P., Cada, V., Martolos, J., Stastny, J., Beran, D., Kolovsky, F., and Kozhukh, D.: Open transport map — Routable OpenStreetMap, in: 2016 IST-Africa Week Conference, pp. 1–11, IEEE, Durban, South Africa, https://doi.org/10.1109/ISTAFRICA.2016.7530657, 2016.

Kadaverugu, R., Sharma, A., Matli, C., and Biniwale, R.: High resolution urban air quality modeling by coupling CFD and mesoscale models: A review, Asia-Pacific Journal of Atmospheric Sciences, 55, 539–556, 2019.

Karl, M., Ramacher, M. O. P., Oppo, S., Lanzi, L., Majamäki, E., Jalkanen, J.-P., Lanzafame, G. M., Temime-Roussel, B., Le Berre, L., and D'Anna, B.: Measurement and Modeling of Ship-Related Ultrafine Particles and Secondary Organic Aerosols in a Mediterranean Port City, 11, 771, https://doi.org/10.3390/toxics11090771, 2023.

Karl, M., Acksen, S., Chaudhary, R., and Ramacher, M. O. P.: Forecasting system for urban air quality with automatic correction and web service for public dissemination, 17, 1–22, https://doi.org/10.1080/17538947.2024.2359569, 2024.

Kelly, B.: Google distance matrix API using OSM, https://github.com/BlaiseKelly/google_speeds, 2025.

Khomenko, S., Pisoni, E., Thunis, P., Bessagnet, B., Cirach, M., Iungman, T., Barboza, E. P., Khreis, H., Mueller, N., Tonne, C., de Hoogh, K., Hoek, G., Chowdhury, S., Lelieveld, J., and Nieuwenhuijsen, M.: Spatial and sector-specific contributions of emissions to ambient air pollution and mortality in European cities: a health impact assessment, The Lancet Public Health, 8, e546–e558, https://doi.org/10.1016/S2468-2667(23)00106-8, 2023.

Kuenen, J., Dellaert, S., Visschedijk, A., Jalkanen, J.-P., Super, I., and Denier Van Der Gon, H.: CAMS-REG-v4: a state-of-the-art high-resolution European emission inventory for air quality modelling, Earth System Science Data, 14, 491–515, https://doi.org/10.5194/essd-14-491-2022, 2022.

Kuik, F., Kerschbaumer, A., Lauer, A., Lupascu, A., Von Schneidemesser, E., and Butler, T. M.: Top–down quantification of NO_x emissions from traffic in an urban area using a high-resolution regional atmospheric chemistry model, 18, 8203–8225, https://doi.org/10.5194/acp-18-8203-2018, 2018.

Li, M., Yu, L., Zhai, Z., He, W., and Song, G.: Development of emission factors for an urban road network based on speed distributions, Journal of Transportation Engineering, 142, 04016 036, 2016.

Li, X., Gu, D., Hohenberger, T. L., Fung, Y. H., Fung, J. C., Lau, A. K., and Liang, Z.: Dynamic quantification of on-road emissions in Hong Kong: impact from traffic congestion and fleet composition variation, Atmospheric Environment, 313, 120 059, 2023.

Macbeth, A. G. and Ltd, V.: A National Road Hierarchy – Are We Ready?, 2007.

Marongiu, A., Distefano, G. G., Moretti, M., Petrosino, F., Fossati, G., Collalto, A. G., and Angelino, E.: Machine Learning Approach for Local Atmospheric Emission Predictions, 2, 380–401, https://doi.org/10.3390/air2040022, 2024.

Moradi, M., Roche, S., and Mostafavi, M. A.: Exploring five indicators for the quality of OpenStreetMap road networks: a case study of Québec, Canada, 75, 1–31, https://doi.org/10.1139/geomat-2021-0012, 2021.

National Atmospheric Emissions Inventory: The UK National Atmospheric Emissions Inventory (NAEI) | National Atmospheric Emissions Inventory, http://naei.energysecurity.gov.uk/, 2025.

Navarro-Barboza, H., Pandolfi, M., Guevara, M., Enciso, S., Tena, C., Via, M., Yus-Díez, J., Reche, C., Pérez, N., Alastuey, A., Querol, X., and Jorba, O.: Uncertainties in source allocation of carbonaceous aerosols in a Mediterranean region, 183, 108 252, https://doi.org/10.1016/j.envint.2023.108252, 2024.

Nieuwenhuijsen, M. J.: Urban and transport planning, environmental exposures and health-new concepts, methods and tools to improve health in cities, Environmental Health, 15, S38, https://doi.org/10.1186/s12940-016-0108-1, 2016.

Ntziachristos, L., Gkatzoflias, D., Kouridis, C., and Samaras, Z.: COPERT: a European road transport emission inventory model, in: Information Technologies in Environmental Engineering: Proceedings of the 4th International ICSC Symposium Thessaloniki, Greece, May 28-29, 2009, pp. 491–504, Springer, 2009.

OpenStreetMap contributors: Open Street Map, published: https://www.openstreetmap.org, 2023.

Pesaresi, M., Schiavina, M., Politis, P., Freire, S., Krasnodębska, K., Uhl, J. H., Carioli, A., Corbane, C., Dijkstra, L., Florio, P., Friedrich, H. K., Gao, J., Leyk, S., Lu, L., Maffenini, L., Mari-Rivero, I., Melchiorri, M., Syrris, V., Van Den Hoek, J., and Kemper, T.: Advances on the Global Human Settlement Layer by joint assessment of Earth Observation and population survey data, International Journal of Digital Earth, 17, 2390 454, https://doi.org/10.1080/17538947.2024.2390454, 2024.

Pulugurtha, S. S. and Mathew, S.: Modeling AADT on local functionally classified roads using land use, road density, and nearest nonlocal road data, Journal of Transport Geography, 93, 103 071, https://doi.org/10.1016/j.jtrangeo.2021.103071, 2021.

Pültz, J., Thürkow, M., Banzhaf, S., and Schaap, M.: Nitrogen Dioxide Source Attribution for Urban and Regional Background Locations Across Germany, 16, 312, https://doi.org/10.3390/atmos16030312, 2025.

Ramacher, M. O. P., Kakouri, A., Speyer, O., Feldner, J., Karl, M., Timmermans, R., Denier Van Der Gon, H., Kuenen, J., Gerasopoulos, E., and Athanasopoulou, E.: The UrbEm Hybrid Method to Derive High-Resolution Emissions for City-Scale Air Quality Modeling, Atmosphere, 12, 1404, https://doi.org/10.3390/atmos12111404, 2021.

Ramacher, M. O. P., Badeke, R., Fink, L., Quante, M., Karl, M., Oppo, S., Lenartz, F., Dury, M., and Matthias, V.: Assessing the effects of significant activity changes on urban-scale air quality across three European cities, 22, 100 264, https://doi.org/10.1016/j.aeaoa.2024.100264, 2025.

Razali, N. A. M., Shamsaimon, N., Ishak, K. K., Ramli, S., Amran, M. F. M., and Sukardi, S.: Gap, techniques and evaluation: traffic flow prediction using machine learning and deep learning, Journal of Big Data, 8, 152, https://doi.org/10.1186/s40537-021-00542-7, 2021.

Samland, M., Badeke, R., Grawe, D., and Matthias, V.: Variability of aerosol particle concentrations from tyre and brake wear emissions in an urban area, 24, 100 304, https://doi.org/10.1016/j.aeaoa.2024.100304, 2024.

Santiago, J. L., Borge, R., Sanchez, B., Quaassdorff, C., Paz, D. d. l., Martilli, A., Rivas, E., and Martín, F.: Estimates of pedestrian exposure to atmospheric pollution using high-resolution modelling in a real traffic hot-spot, Science of The Total Environment, 755, 142 475, https://doi.org/https://doi.org/10.1016/j.scitotenv.2020.142475, 2021.

Sekuła, P., Marković, N., Vander Laan, Z., and Sadabadi, K. F.: Estimating historical hourly traffic volumes via machine learning and vehicle probe data: A Maryland case study, Transportation Research Part C: Emerging Technologies, 97, 147–158, https://doi.org/10.1016/j.trc.2018.10.012, 2018.

Sharma, S., Lingras, P., Xu, F., and Kilburn, P.: Application of neural networks to estimate AADT on low-volume roads, Journal of Transportation Engineering, 127, 426–432, publisher: American Society of Civil Engineers, 2001.

Shen, Y., De Hoogh, K., Schmitz, O., Gulliver, J., Vienneau, D., Vermeulen, R., Hoek, G., and Karssenberg, D.: Europe-wide high-spatial resolution air pollution models are improved by including traffic flow estimates on all roads, Atmospheric Environment, 335, 120 719, https://doi.org/10.1016/j.atmosenv.2024.120719, 2024.

Skoulidou, I., Koukouli, M.-E., Manders, A., Segers, A., Karagkiozidis, D., Gratsea, M., Balis, D., Bais, A., Gerasopoulos, E., Stavrakou, T., Van Geffen, J., Eskes, H., and Richter, A.: Evaluation of the LOTOS-EUROS NO_2 simulations using ground-based measurements and S5P/TROPOMI observations over Greece, 21, 5269–5288, https://doi.org/10.5194/acp-21-5269-2021, 2021.

Smit, R., Smokers, R., and Rabé, E.: A new modelling approach for road traffic emissions: VERSIT+, Transportation Research Part D: Transport and Environment, 12, 414–422, 2007.

Sun, X., Das, S., et al.: Developing a method for estimating AADT on all Louisiana roads., Tech. rep., Louisiana Transportation Research Center, 2015.

Tang, F. and Ishwaran, H.: Random forest missing data algorithms, Statistical Analysis and Data Mining: The ASA Data Science Journal, 10, 363–377, https://doi.org/10.1002/sam.11348, 2017.

Thunis, P., Pisoni, E., Sajani, S., Monforti, F., Bessagnet, B., Vignati, E., and de Meij, A.: Urban PM2.5 Atlas. Air Quality in European Cities - 2023 Report, JRC, https://doi.org/10.2760/63641, 2023.

Treiber, M., Kesting, A., and Thiemann, C.: How Much does Traffic Congestion Increase Fuel Consumption and Emissions? Applying a Fuel Consumption Model to the NGSIM Trajectory Data, Annual Meeting of the Transportation Research Board, 2007.

Trombetti, M., Pisoni, E., Lavalle, C., et al.: Downscaling methodology to produce a high resolution gridded emission inventory to support local/city level air quality policies, Office for Official Publications of the European Communities, Luxembourg EUR, 28428, 2017.

Valencia, V. H., Levin, G., and Ketzel, M.: Downscaling global anthropogenic emissions for high-resolution urban air quality studies, Atmospheric Pollution Research, 13, 101 516, https://doi.org/10.1016/j.apr.2022.101516, 2022.

Venter, Z. S., Figari, H., Krange, O., and Gundersen, V.: Environmental justice in a very green city: Spatial inequality in exposure to urban nature, air pollution and heat in Oslo, Norway, Science of The Total Environment, 858, 160 193, https://doi.org/10.1016/j.scitotenv.2022.160193, 2023.

Wen, Y., Wu, R., Zhou, Z., Zhang, S., Yang, S., Wallington, T. J., Shen, W., Tan, Q., Deng, Y., and Wu, Y.: A data-driven method of traffic emissions mapping with land use random forest models, Applied Energy, 305, 117 916, https://doi.org/10.1016/j.apenergy.2021.117916, 2022.

World Health Organization and others: WHO global air quality guidelines: particulate matter (PM2. 5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide, World Health Organization, 2021.

Xue, H., Jiang, S., and Liang, B.: A Study on the Model of Traffic Flow and Vehicle Exhaust Emission, Mathematical Problems in Engineering, 2013, 1–6, https://doi.org/10.1155/2013/736285, 2013.

Yang, D., Zhang, S., Niu, T., Wang, Y., Xu, H., Zhang, K. M., and Wu, Y.: High-resolution mapping of vehicle emissions of atmospheric pollutants based on large-scale, real-world traffic datasets, Atmospheric Chemistry and Physics, 19, 8831–8843, https://doi.org/10.5194/acp-19-8831-2019, 2019.

Zanaga, D., Van De Kerchove, R., Daems, D., De Keersmaecker, W., Brockmann, C., Kirches, G., Wevers, J., Cartus, O., Santoro, M., Fritz, S., et al.: ESA WorldCover 10 m 2021 v200, 2022.