CAMS Service Evolution

D6.2 Methodology/prototype for uncertainty estimates in the source contribution to EU cities

Due date of deliverable	August 30, 2025
Submission date	9/9/25
File Name	CAMEO-D6.2-V1.1
Work Package /Task	WP6, Task 6.2 + 6.3.3
Organisation Responsible of Deliverable	MET Norway
Author name(s)	Hilde Fagerli, Lewis Blake, Svetlana Tsyro, Eivind Grøtting Wærsted, Alvaro Valdebenito, Peter Wind, Palmira Messina, Elsa Real, Blandine Raux, Ingrid Super, Jeroen Kuenen
Revision number	V1.1
Status	Issued
Dissemination Level	Public

Funded by the European Union

The CAMEO project (grant agreement No 101082125) is funded by the European Union.
Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the Commission. Neither the European Union nor the granting authority can be held responsible for them.

1 Executive Summary

The Copernicus Atmosphere Monitoring Service (CAMS) currently offers a number of products to support decision and policy makers in regard to mitigation of air pollution at the European scale (https://policy.atmosphere.copernicus.eu/). These policy products provide information about the causes and main drivers of air pollution in European cities, so-called source allocation/attribution products, and their potential evolution in the future. Overall, better knowledge of the uncertainties in the sector/source attribution in the CAMS policy products may help policy makers to make more informed decisions when prioritizing measures for specific activity sectors.

We have developed methodologies and uncertainty estimates directly relevant for the CAMS Policy Support Service. More specifically, we have:

- Established uncertainty estimates in CAMS source receptor policy products related to emissions, model resolution and city area definition
- Established the link to urban scale by developing an in-city variability/correction estimate for the local plus non-local contribution

How sensitive is the local contribution (city to itself) to the geographical definition of the city area and the CTM model resolution?

When designing air quality policies, it is important to know how much can be achieved by targeting local emissions (e.g., from a city planning perspective), and to what extent either national plans or international agreements should be prioritized. Therefore it is important to quantify the uncertainty in the local contribution product presented by CAMS.

Our analysis shows that:

- The calculated local contribution is very sensitive to both the model resolution and the geographical definition of the source & receptor regions, as well as to the averaging method employed within the receptor area.
- Area-weighted local contributions (as in the current CAMS Policy Support Service) are systematically lower than population weighted local contributions (as in GAINS) which again are systematically lower than using the center grid cell as the receptor (similar to SHERPA & JRC Urban $PM_{2.5}$ atlas, although they use the maximum concentration grid cell). Therefore, the current CAMS Policy Support Service systematically attributes a smaller responsibility to the cities for its own pollution - for $PM_{2.5}$, PM_{10} and NO_2 than other policy relevant products. For O_3, the titration is highest when using the centre grid receptor and lowest for area weighted receptors.
- Higher model resolution increases the modelled local contribution systematically, by a factor 2-3 for $PM_{2.5}$, PM_{10} and NO_2 (when increasing resolution from 0.3x0.2 degree to 0.2x0.1 degree). For ozone, the titration by local city emissions is halved. The resolution of the current CTM models running for CAMS Policy Service was changed from coarser resolutions (different in each model) to 0.2x0.1 degree from 1.1.2025, partly due to the results discussed here. Note, however, that the total concentrations (local plus non-local) are not consistently different due to resolution, except for ozone, where annual ozone is systematically lower in cities with high titration.

From a policy perspective it is not clear which definitions of source and receptor areas should be used. On one hand one might argue that using the centre grid (or the grid with maximum concentration) as receptor is most relevant, as this would ensure that the sources which are

most important for the worst areas are targeted when planning mitigation actions - in line with achieving the limit values set in the Ambient Air Quality Directive. On the other hand, using the population weighted concentrations would be most relevant when policy is acting on reducing health effects overall. For the source area, it seems most reasonable to choose the core city area (the administrative area), as this is the area that city authorities can act on. In the end, the decisions should be made by the users, and the results discussed here are planned to be presented and discussed at several meetings and workshops during fall and winter, such as FAIRMODE and the next CAMS Policy User workshop.

Link between regional and urban scales

There have been repeated user requests in the CAMS Policy User Workshops that CAMS should deliver air quality information at finer spatial resolution, making the results more relevant for the urban scale. Linking regional and urban scales in the CAMS policy products will make these products more relevant for policy analysis such as the impact assessments being done at present for the Ambient Air Quality Directive, where exceedances in hot spot areas possibly will be driving policy.

In the work undertaken for this deliverable, linkages between urban and regional scales have been made using a coupled regional and urban-scale modelling system (EMEP/uEMEP). First, we analysed the impact of resolution by comparing the coarser resolution results in the present source receptor service with the EMEP model (0.2x0.1 degrees) to fine resolution results from uEMEP at 250m for NO_2 and $PM_{2.5}$ for all the cities in the CAMS Policy Support Service. For the winter period of 2019, population-weighted mean NO_2 concentrations increase on average by 40 %, but it varies importantly between cities, with some cities having close to no increase at all and others more than doubling. In most cities, the concentrations calculated at observational sites with uEMEP are closer to the observed mean, although the NO_2 concentration is still underestimated on average after downscaling. Also, the spatial correlation between the stations is strongly improved by the downscaling for nearly all the cities, and on average it is improved from 0.26 to 0.64 (0.71) when using 250 m (50 m) resolution and compared to EEA stations.

Furthermore, downscaling with uEMEP to 250 m can give important additional information on exposure, both by correcting the average population exposure and by accounting for the variability within the core city that gives rise to higher exposure for a subset of the population. For instance, the limit value for daily mean NO_2 concentration (50 µg/m3) in the new European Ambient Air Quality Directive (AAQD) is exceeded much more frequently after downscaling, both in terms of the population-weighted mean (238% more days) and for the 95% percentile (219% more days).

The relative increase in total $PM_{2.5}$ concentration due to downscaling is smaller than for NO_2 for most cities, on average a 18 % increase. The city-mean bias is improved from -2.7 to -0.9 (-0.6) µg/m3 for uEMEP with 250 m (50 m) resolution, but the spatial correlation is not systematically improved across the cities. Downscaling mainly has an impact on population-weighted $PM_{2.5}$ in only a smaller number of cities where residential heating has a very large contribution. The reason for the smaller improvement of $PM_{2.5}$ is most probably related to the quality of the gridded primary PM emissions (and specifically residential heating) as well as the proxy used to downscale those emissions at European scale. It has been shown that when using proxy data that are more closely linked to residential combustion, as available for Norway, the downscaling can give significant improvements.

The number of days above the new AAQD limit value for daily mean $PM_{2.5}$ (25 µg/m3) increases by (only) 31% for the city mean and by 52% for the 95^{th} percentile. However, the very high values of daily mean $PM_{2.5}$ above 50 µg/m3 (and even 100 µg/m3) occur much more frequently, particularly in the 95^{th} percentile. Since the downscaling of $PM_{2.5}$ does not

systematically improve performance, this increase in occurrence of very high concentrations might be less reliable.

Secondly, we set out to develop an operational in-city variability/correction estimate that could account for the resolution issue (e.g. represent the effect of downscaling EMEP model results by uEMEP). Our analysis demonstrates that there is a correlation between the stability function and the variability in increase in traffic NO_x contribution when downscaled by uEMEP. However, it is clear that stability alone cannot reproduce the time variation in the relative increase in NO_x traffic contribution, and the relationship with stability is more of qualitative nature than a simple one-to-one dependency. This is not surprising, since other factors such as boundary layer height, wind speed and wind direction will also affect the downscaling. Due to the important time-variation in the effect of downscaling, it seems overly simplistic to apply a constant correction factor for each city at all times in the CAMS Policy Support service.

However, it is manageable to run uEMEP downscaling for all the 80 cities included in CAMS Policy Service for the 4-day forecast horizon operationally. Based on the computation times of the runs of this study, all cities can be downscaled (to 250m) within 1 hour on a computational node with 96 parallel CPUs, each with 2 GB memory. If a coarser resolution of 500 m or 1 km was used instead, the computational cost would be lower.

Propagation of emission uncertainties into CAMS Policy Products

The largest single source of uncertainty in air pollution forecasting (and source apportionment) is likely related to emissions. Although there have been several studies where emission uncertainties have been propagated into air quality forecasts/source apportionment, to our knowledge there has been no attempt to generate such information in near-real-time (NRT). Clearly, such information will be very valuable for interpreting the drivers behind episode situations, during the episode itself, supporting communication both to the public as well as to policy makers.

Emission uncertainties have been developed in WP5 of CAMEO: uncertainties in total emissions per country and sector for NO_x, SO_x, NH_3, NMVOC and PPM, uncertainties in gridded PPM emissions and uncertainties in the temporal disaggregation of emissions (per country, sector and component).

Propagation of emission uncertainties have been investigated using two different CAMS Policy Support source allocation products for cities: i) the potential impact from country and city emissions and ii) the sector apportionment from the ACT metamodel. In principle, a similar analytic approach (tailored for CAMS operational applications) has been used for both of them, although there are some important differences.

City & country contributions to City SR

The propagation of emission uncertainties for the city and country to city SR service is done analytically for $PM_{2.5}$. We take advantage of the local fraction (LF) calculations done with the EMEP MSC-W model, where all the derivatives (the change in $PM_{2.5}$ due to changes in the different emissions per country, sector and component) are calculated. In addition, the city-to-itself contribution is derived. Normally, the number of model runs that would have been needed if this should be done with Brute Force calculations, reducing one country, sector, and component at a time (30 (countries) x 5 (components) x 13 (sectors) = 1950 model runs) would be too large for this to be possible in an operational setting with time restrictions, but with LF this is computationally feasible . The uncertainties can be calculated by accounting for the uncertainties attributed to each of the contributions (from each country, sector, component plus the local contribution) to anthropogenic $PM_{2.5}$ concentrations, plus the covariance terms (or the correlation in uncertainties between all the terms).

We have explored different correlation structures for the uncertainties (uncorrelated, correlated across the same sector in different countries and components, and fully correlated). In general, the correlation structure is important, with the uncertainties in the correlated case being a factor 2 higher than the uncorrelated case.

Further, we have added the uncertainty related to the temporal breakdown of yearly emissions (per country and sector) by monthly, day-of-week, and hour-of-day time factors. Since we can not assume that the temporal uncertainty at time t manifests instantaneously into the concentration uncertainty at time t, different forms of aggregation methods have been investigated. Generally we find that the 24-hour rolling window, mean, and median yield relatively similar results. On the other hand, aggregating the maximum uncertainty results in the most noticeable change in the total prediction interval.

We have evaluated the method's performance across selected air pollution episodes in Europe during 2019. Overall, the approach produces plausible uncertainty estimates, offering potential value to users of CAMS Policy data products by quantifying the uncertainty associated. However, there are important limitations to keep in mind. Most critically, the underlying emissions' uncertainties themselves are highly uncertain, often due to limited data, expert judgment, or assumptions in the original inventories. This adds a layer of ambiguity to any propagated uncertainty estimates. Second, the degree of correlation between emissions could significantly influence the results, but in practice, the true correlation structure is poorly understood and rarely quantified.

Leveraging on the metamodel ACT

A similar analytic methodology for propagating emission uncertainties to $PM_{2.5}$ concentrations using the ACT metamodel has been explored, accounting for both uncertainties in temporal allocation factors and total annual emissions.

The current implementation provides estimates of variability in $PM_{2.5}$ concentrations over major European cities, but the methodology can be readily extended to all grid points within the modeling domain and to additional pollutants. The method is computationally efficient, requiring minimal processing time and resources, making it particularly well suited for potential operational deployment.

Results indicate relatively large propagated errors, which are consistent with the substantial uncertainties associated with emissions—especially those related to temporal disaggregation factors. The residential sector is by far the most affected by temporal uncertainties, and consequently, the total propagated errors are highest in regions and periods where this sector dominates.

Notable differences arise when comparing the results based on ACT to the method leveraging on LF in the EMEP MSC-W model. These discrepancies can be attributed to several factors. While both methods adopt an analytical approach to variance estimation, the results based on LF in the EMEP MSC-W model also account for some degree of correlation between emission sources. Further, ACT simplifies the uncertainties taken into account to only be a function of emission sectors, while the EMEP MSC-W model results takes into account uncertainties per country, sector and component. Finally, the two underlying models, ACT and the Local Fraction-based method, differ in the underlying model results for $PM_{2.5}$ and the contributions themselves.

Table of Contents

2 Introduction

2.1 Background

Monitoring the composition of the atmosphere is a key objective of the European Union's flagship Space programme Copernicus, with the Copernicus Atmosphere Monitoring Service (CAMS) providing free and continuous data and information on atmospheric composition.

The CAMS Service Evolution (CAMEO) project will enhance the quality and efficiency of the CAMS service and help CAMS to better respond to policy needs such as air pollution and greenhouse gases monitoring, the fulfilment of sustainable development goals, and sustainable and clean energy.

CAMEO will help prepare CAMS for the uptake of forthcoming satellite data, including Sentinel-4, -5 and 3MI, and advance the aerosol and trace gas data assimilation methods and inversion capacity of the global and regional CAMS production systems.

CAMEO will develop methods to provide uncertainty information about CAMS products, in particular for emissions, policy, solar radiation and deposition products in response to prominent requests from current CAMS users.

CAMEO will contribute to the medium- to long-term evolution of the CAMS production systems and products.

The transfer of developments from CAMEO into subsequent improvements of CAMS operational service elements is a main driver for the project and is the main pathway to impact for CAMEO.

The CAMEO consortium, led by ECMWF, the entity entrusted to operate CAMS, includes several CAMS partners thus allowing CAMEO developments to be carried out directly within the CAMS production systems and facilitating the transition of CAMEO results to future upgrades of the CAMS service.

This will maximise the impact and outcomes of CAMEO as it can make full use of the existing CAMS infrastructure for data sharing, data delivery and communication, thus supporting policymakers, business and citizens with enhanced atmospheric environmental information.

2.2 Scope of this deliverable

2.2.1 Objectives of this deliverables

The work described in this deliverable has two objectives:

- Establish uncertainty estimates in CAMS source receptor policy products related to emissions, model resolution and city area definition
- Generate the link to urban scale by developing a variability/correction estimate for the local contribution

2.2.2 Work performed in this deliverable

In this deliverable the work as planned in the Description of Action (DoA, WP6 T6.2 and T6.3.3) was performed.

2.2.3 Deviations and counter measures

No deviations have been encountered.

2.2.4 CAMEO Project Partners:

ECMWF	EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS
Met Norway	METEOROLOGISK INSTITUTT
BSC	BARCELONA SUPERCOMPUTING CENTER-CENTRO NACIONAL DE SUPERCOMPUTACION
KNMI	KONINKLIJK NEDERLANDS METEOROLOGISCH INSTITUUT-KNMi
SMHI	SVERIGES METEOROLOGISKA OCH HYDROLOGISKA INSTITUT
BIRA-IASB	INSTITUT ROYAL D'AERONOMIE SPATIALEDE BELGIQUE
HYGEOS	HYGEOS SARL
FMI	ILMATIETEEN LAITOS
DLR	DEUTSCHES ZENTRUM FUR LUFT - UND RAUMFAHRT EV
ARMINES	ASSOCIATION POUR LA RECHERCHE ET LE DEVELOPPEMENT DES METHODES ET PROCESSUS INDUSTRIELS
CNRS	CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE CNRS
GRASP-SAS	GENERALIZED RETRIEVAL OF ATMOSPHERE AND SURFACE PROPERTIES EN ABREGE GRASP
CU	UNIVERZITA KARLOVA
CEA	COMMISSARIAT A L ENERGIE ATOMIQUE ET AUX ENERGIES ALTERNATIVES
MF	METEO-FRANCE
TNO	NEDERLANDSE ORGANISATIE VOOR TOEGEPAST NATUURWETENSCHAPPELIJK ONDERZOEK TNO
INERIS	INSTITUT NATIONAL DE L ENVIRONNEMENT INDUSTRIEL ET DES RISQUES - INERIS
IOS-PIB	INSTYTUT OCHRONY SRODOWISKA - PANSTWOWY INSTYTUT BADAWCZY
FZJ	FORSCHUNGSZENTRUM JULICH GMBH
AU	AARHUS UNIVERSITET
ENEA	AGENZIA NAZIONALE PER LE NUOVE TECNOLOGIE, L'ENERGIA E LO SVILUPPO ECONOMICO SOSTENIBILE

3 Impact of geographical definition of city as source and receptor and recommendations for use in CAMS

3.1 Introduction

It is well known that the choices made in terms of geographical definition of sources and receptor areas (and indicators) for the cities have a large impact on the results in terms of how much of the pollution that the city itself is responsible for (Thunis et al., 2021).

At present, the geographical city in the CAMS Policy support service is defined as the area corresponding to the 3x3 grid cells in 0.25x0.125 degree. Previously, the CTM models were run at this resolution, and although the models are now run on 0.2x0.1 degree for the CAMS Policy support service(twice the resolution as what is run in the CAMS regional forecast), the geographical definition is exactly the same (hereafter referred to as 3x3). The definition of the source area (the area for which the emissions are reduced or labelled) is the same as the receptor area (the area for which concentrations are averaged). The results on the CAMS Policy support service are presented as the area weighted concentrations in the city receptor.

Several other definitions of city source and receptor areas are in use, for instance in the work done with GAINS for the Clean Air Outlook for EU (CAO4, 2025) or the Gothenburg Protocol, or with the SHERPA modelling tool (e.g. Thunis et al., 2021), see Table 4.1.

In SHERPA (Thunis et al, 2018), the city source area definition is mostly FUA (Functional Urban Area). Thunis et al, 2018 explains that FUAs encompass both the city core and its commuting zone, providing a more comprehensive representation of urban emissions and their impacts on air quality. However, one might argue that using the core city as the source area is more relevant, as this is the area that city authorities have the power to act on, and measures can only be applied here.

It has already been demonstrated in Thunis et al, 2021 that using FUA as source area gives a city contribution or responsibility to $PM_{2.5}$ levels (yearly levels) that is approximately twice as large as using core city as the source. In both cases the grid with the maximum concentration was chosen as the receptor, thus the impact from a larger area (FUA) is by definition larger than the impact of a smaller area (city core) to the same receptor.

In our work we have investigated the sensitivity to different definitions of source and receptor areas of the total contribution as well as the city contribution (not only the max or centre grid cell). Furthermore, we extended the work to cover all of the species for which results are shown in the CAMS Policy Support Service: $PM_{2.5,}$ PM_{10}, NO_2 and O_3.

The main goal of this work has been to demonstrate how sensitive our results are to these choices and to recommend how this should be done in CAMS2_71. We have also investigated the impact of using different resolutions in the CAMS Policy Support Service. The results and recommendations from the work on resolution have already partly been used as basis for changes implemented in CAMS2_71 recently (i.e. all the models in the CAMS Policy Support Service are now running on 0.2x0.1 degree resolution).

Table 3.1 Overview of source and receptor definitions used in the CAMS Policy Support Service, JRC Urban Atlas (Thunis et al., 2017) and GAINS (Amann et al., 2011)

	City source definition	City receptor definition
CAMS Policy Support	3x3 grids (0.25x0.125 degree)	3x3 grids (0.25x0.125), area weighted concentration mean
JRC PM2.5 Urban Atlas	City Core, FUA, Urban Audit, 2021	Grid with maximum concentration
GAINS	City Core, FUA (GHS Urban Centre Database, Florczyk et al., 2019)	Population weighted concentration in core city (but only primary PM is considered for the city contribution to itself)

3.2 Experimental setup

The EMEP MSC-W model version rv5.0 has been used for work in this deliverable. The horizontal resolution is 0.2x0.1, with 20 vertical layers (the lowest with a height of approximately 50 meters), unless otherwise stated.

Meteorology, emissions, boundary conditions and forest fires for 2019 have been used as input. Boundary conditions come from the CAMS global reanalysis (EAC4) for 2019 (doi.org/10.24381/d58bbf47, Iness et al., 2019), read every 3 hours. Meteorological data is taken from the ECMWF-IFS operational 12 UTC meteorological forecasts for 2019 (ECMWF-IFS 45r1 until June 11 and 46r1 from June 12th). Anthropogenic emissions are taken from the CAMS REG AP v6.1 emission data set. The forest fire emissions are taken from GFAS v1.2 (Kaiser et al., 2012, doi.org/10.24381/a05253c7), while the model also includes climatological ocean DMS and soil-NOx emissions. For temporal distribution of the emissions for the year 2019, the CAMS-REG-TEMPO v4.1 was used.

All anthropogenic emissions were reduced city by city (by 15%, all emissions at the same time) for all the 79 cities shown on the CAMS Policy User interface. The difference between the base run (full emissions) and the city reduction runs were used to quantify the city contribution by multiplying by 100/15. Other anthropogenic contributions were calculated by a model run where all anthropogenic emissions were reduced by 15%, and the city contribution was removed. The 'rest' contribution is defined as the base run minus the anthropogenic contributions.

We have investigated using 3 different definitions for the source area:

- Corresponding to 3x3 grid cells in 0.25x0.125 degree (the setup used in CAMS Policy Support Service when this work started)
- Core city (Urban audit, 2021) - also used in the JRC Urban Atlas for $PM_{2.5}$
- Functional Urban area (Urban Audit, 2021) - also used in the JRC Urban Atlas for $PM_{2.5}$)

Furthermore, we analyzed results when using the following definitions of the receptor area (where core city area or FUA are chosen as source areas):

- Area weighted core city
- Area weighted FUA

- Population weighted core city
- Population weighted FUA
- 1-grid (the geographical mid-point)

The population data is Global Human Settlement Layer (GHSL) 2020, described in Pesaresi et al, 2024. In addition, model runs with a selection of the definitions above have been run with a resolution of 0.3x0.2 in order to investigate the impact of resolution.

3.3 Results

3.3.1 Differences due to source area definition

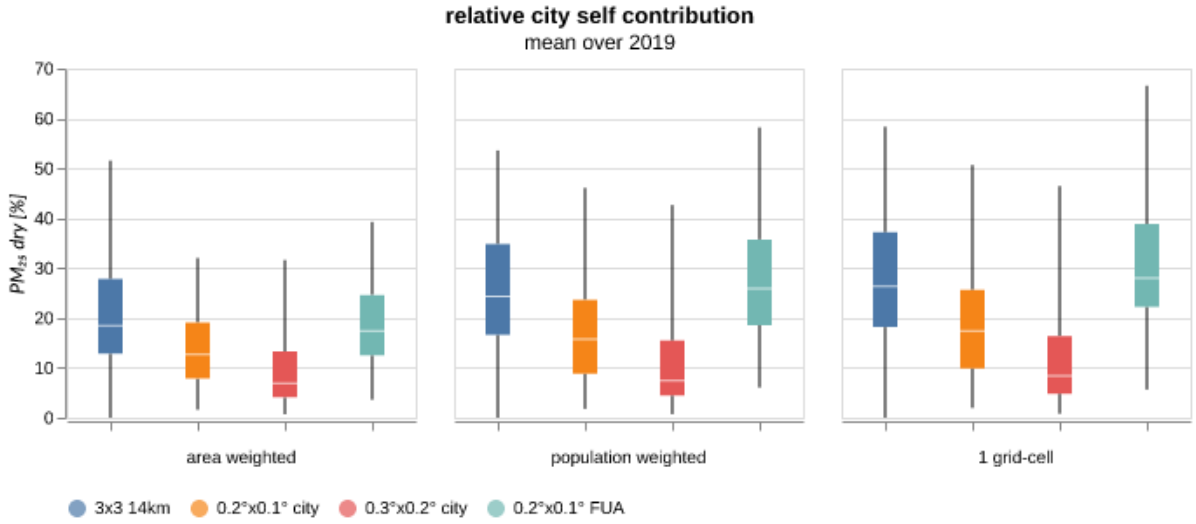
In SHERPA (Thunis et al, 2018), the city source area definition is mostly FUA. Thunis et al, 2018 explains that FUAs encompass both the city core and its commuting zone, providing a more comprehensive representation of urban emissions and their impacts on air quality.

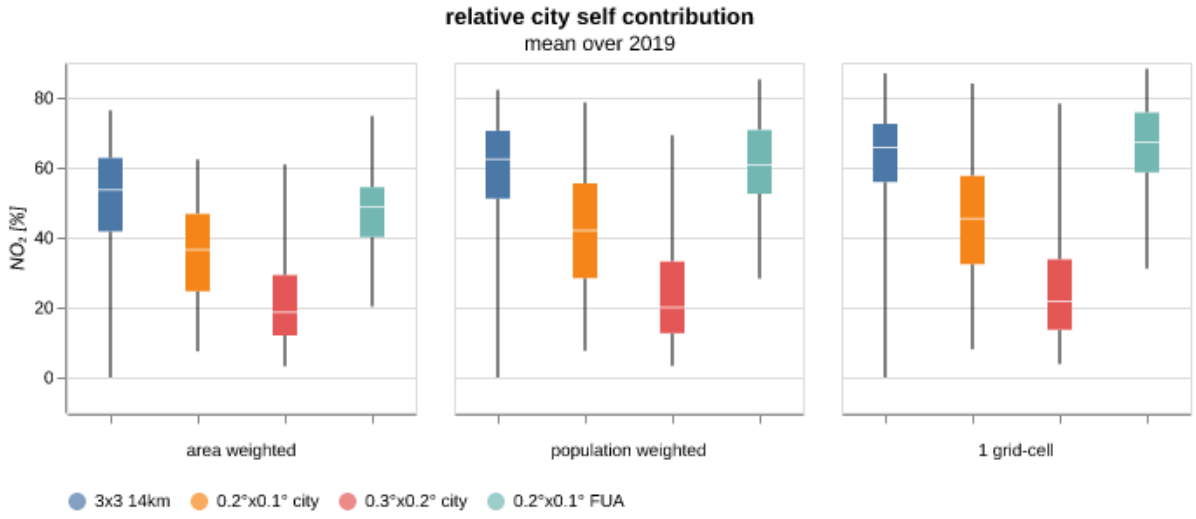
One might argue that using the core city as the source area is more relevant, as this is the area that city authorities have the power to act on, and measures can only be applied here.

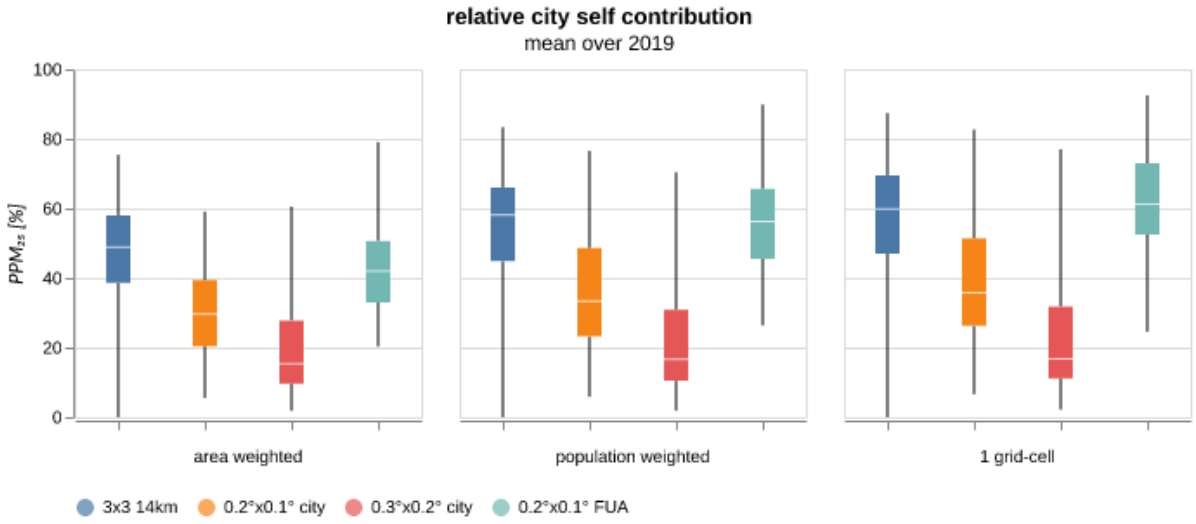
It has already been demonstrated in Thunis et al, 2021 that using FUA as source area gives a city contribution or responsibility to $PM_{2.5}$ levels (yearly levels) that is approximately twice as large as using core city as the source. In both cases the grid with the maximum concentration was chosen as the receptor, thus the impact from a larger area (FUA) is by definition larger than the impact of a smaller area (city core) to the same receptor.

In our work we have looked at different city source definitions, but we also apply other receptor definitions (not only the max or centre grid cell). Furthermore, we extended the work to cover all of the species for which results are shown in the CAMS Policy Support Service: $PM_{2.5,}$ PM_{10}, NO_2 and O_3.

Figure 3.1 shows box quantile diagrams summarizing the city contribution to $PM_{2.5}$ (top) for the 79 cities investigated. Similar to Thunis et al, 2021, we find that the city contribution to $PM_{2.5}$ is about 30% when the source area is FUA and the centre grid is receptor (assuming that in most cases the centre grid would be equal to the grid with highest concentrations). This is reduced to a bit less than 20% when the core city is used as a source area. For the source definition used in CAMS2_71 (3x3 grid), the relative city self contribution is more similar to using FUA (a mean of around 27%), probably because the 3x3 grid area is more similar in size to FUA than to the core city. When the source area and the receptor area are the same, the core city definition systematically gives lower local fraction - between 12-15 percent depending on whether results are area weighted or population weighted.


relative city self contribution
mean over 2019
$PM_{2.5}$ dry [%]
70
60
50
40
30
20
10
0
area weighted
population weighted
1 grid-cell
3x3 14km
0.2°x0.1° city
0.3°x0.2° city
0.2°x0.1° FUA


relative city self contribution
mean over 2019
NO_2 [%]
80
60
40
20
0
area weighted
population weighted
1 grid-cell
3x3 14km
0.2°x0.1° city
0.3°x0.2° city
0.2°x0.1° FUA


relative city self contribution
mean over 2019
$PPM_{2.5}$ [%]
100
80
60
40
20
0
area weighted
population weighted
1 grid-cell
3x3 14km
0.2°x0.1° city
0.3°x0.2° city
0.2°x0.1° FUA

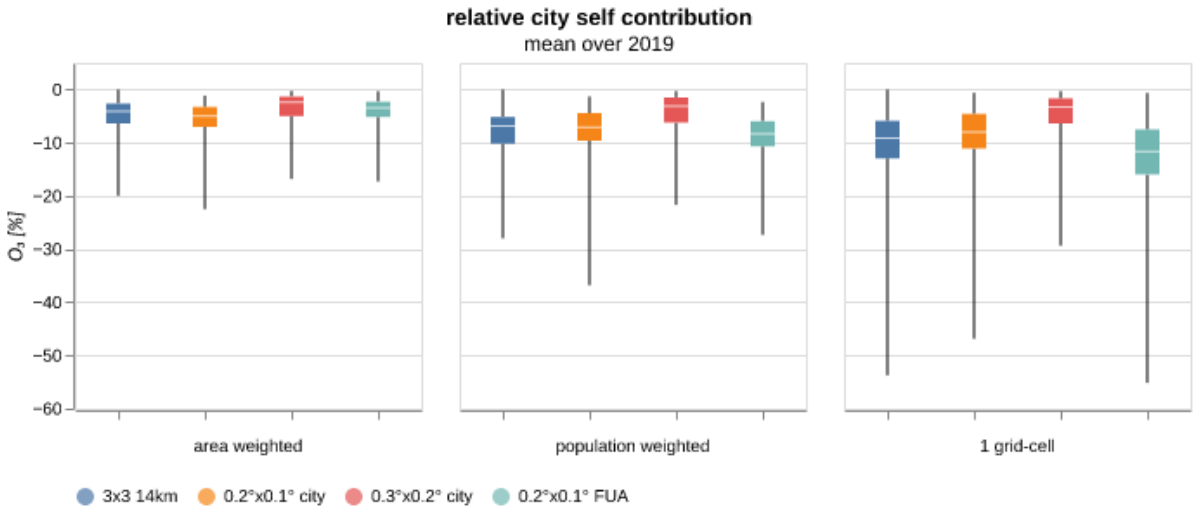


Figure 3.1 Box quantile diagrams summarizing the local contribution to $PM_{2.5}$, NO_2 , primary $PM_{2.5}$ and O_3. concentrations for the 79 European cities, For each of the 3 plots, the receptor is the same: from left: 1) area weighted receptor = source, 2) population weighted receptor = source and 3) centre grid cell. Within each plot, sources (and model resolution) vary: '3x3 14 km' means that the city source area is defined by 3x3 grid cells in 0.25x0.125 degree model resolution, '0.2x0.1 city' means that core city area is the source and the model is run in 0.2x0.1 degree resolution, '0.3x0.2 city' means that core city area is the source and the model is run in 0.3x0.2 degree resolution and '0.2x0.1 FUA' means that FUA is the source area and the model is run in 0.2x0.1 degree resolution. The box borders show the 10th and 90th percentile and the two extremities are the max and min values. The horizontal line within the box represents the median.

For NO_2 (Figure 3.1, second row), the contribution from the city itself is substantially higher. When the city is defined by FUA, and the centre grid is the receptor, the median yearly contribution (of the 79 cities) is around 70%. For NO_2 as for $PM_{2.5}$, the results for the 3x3 grid are very similar to those of FUA to the centre grid. Using core city as the source area brings the city responsibility down to around 40%, and lowest for the area weighted receptor.

For reference, we also show the results for primary $PM_{2.5}$ (third row). Since this is a primary species, it shows very similar results as NO_2 in terms of the local contribution from the city, e.g. 60% in the centre grid cell comes from FUA (against 70% for $NO_{2)}$ and 35% from core city (against 47% for NO_2).

For ozone (bottom row, Figure 3.1), the city contribution is always negative. FUA has the most negative impact in general, except for the area weighted results where the core city to core city is most negative. For the population weighted receptor results and for the centre grid receptor results, the contribution from the city is around - 8-12% (except for the coarser resolution runs, where titration is smaller).

The appendix includes also figures for PM_{10}, primary PM_{10} and the secondary species SO_4, NO_3 and NH_4.

In Figure 3.2 we show how the city contribution varies with time when using the core city (upper row) or FUA (bottom row) as source (example for Amsterdam March 2019). The receptor is either the same as the source (area weighted or population weighted) or the centre grid cell. Typically, the local contribution is higher for FUA than for core city, although in these plots it is mostly clear for the plots where the receptor is the centre grid cell.

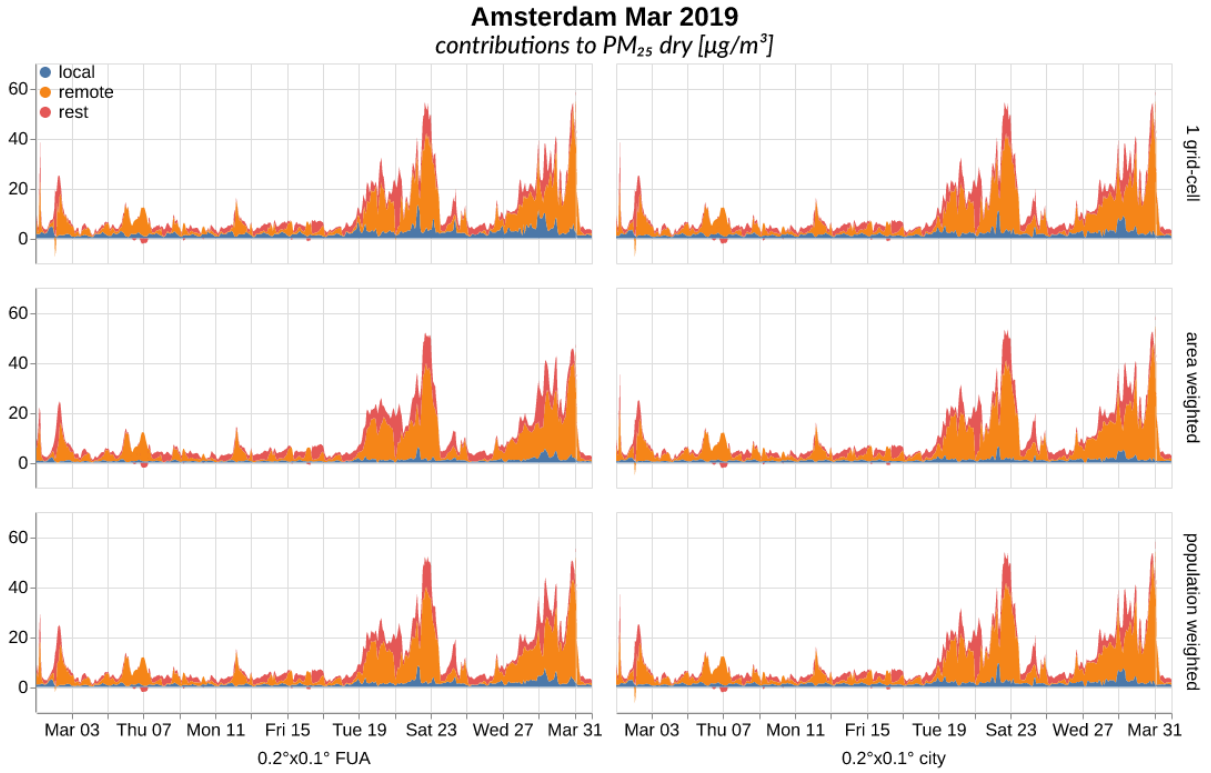


Figure 3.2 Contribution from city, remote (all anthropogenic emissions outside the city) and rest (total minus all sources). From top to bottom: receptor area 1-grid-cell, area weighted, population weighted. Left 3 plots: Core city source area. Right 3 plots: FUA source area

3.3.2 Differences due to receptor area definition

In the current CAMS Policy Support Service, the geographical definition of source and receptor is the same, and the results are provided as area weighted means.

Thunis et al, 2021 argues that using the grid cell with highest concentration (the 'worst case receptor') as the receptor ensures that the model targets the exceedance zones most effectively, in line with EU Directive 2008/50/EC. In our work we have used the centre grid cell instead of the grid with the maximum concentration, since we were also looking at finer time resolution (e.g. daily), and it was impractical to sample the maximum concentration in grid when the selected grid was varying between time steps. However, with the resolution here (0.2x0.1 degree), the centre grid and the grid with maximum concentration is in most cases the same.

Figure 3.3 compares results for local contribution to daily, monthly and yearly concentrations of $PM_{2.5}$, using either the centre grid cell as receptor, the area weighted mean or the population weighted mean of the core city (the source is always core city).
Clearly, local contributions are highest for the receptor being the centre grid cell and lowest for the area weighted mean, with the population weighted mean in between. This can also be seen from Figure 3.1, where FUA to the centre grid has a median local contribution of 30%, FUA to population weighted FUA 27% and FUA to area weighted FUA 18%. Similar numbers for core city as source is 18% (centre grid receptor), 16% (population weighted core city) and 13% (area weighted core city).

The box quantile diagram for NO_2 (and $PPM_{2.5}$) shows a similar picture, where the local fraction for FUA, core city or 3x3 to the receptor always follow centre grid > population

weighted> area weighted for the receptor.
For ozone, the centre grid as a receptor always gives the largest negative impact of the city, with the population weighted results being quite similar. Area weighted receptor areas in general give smaller negative impacts of the city, probably because population density is highest near roads where NO_x emissions are highest.

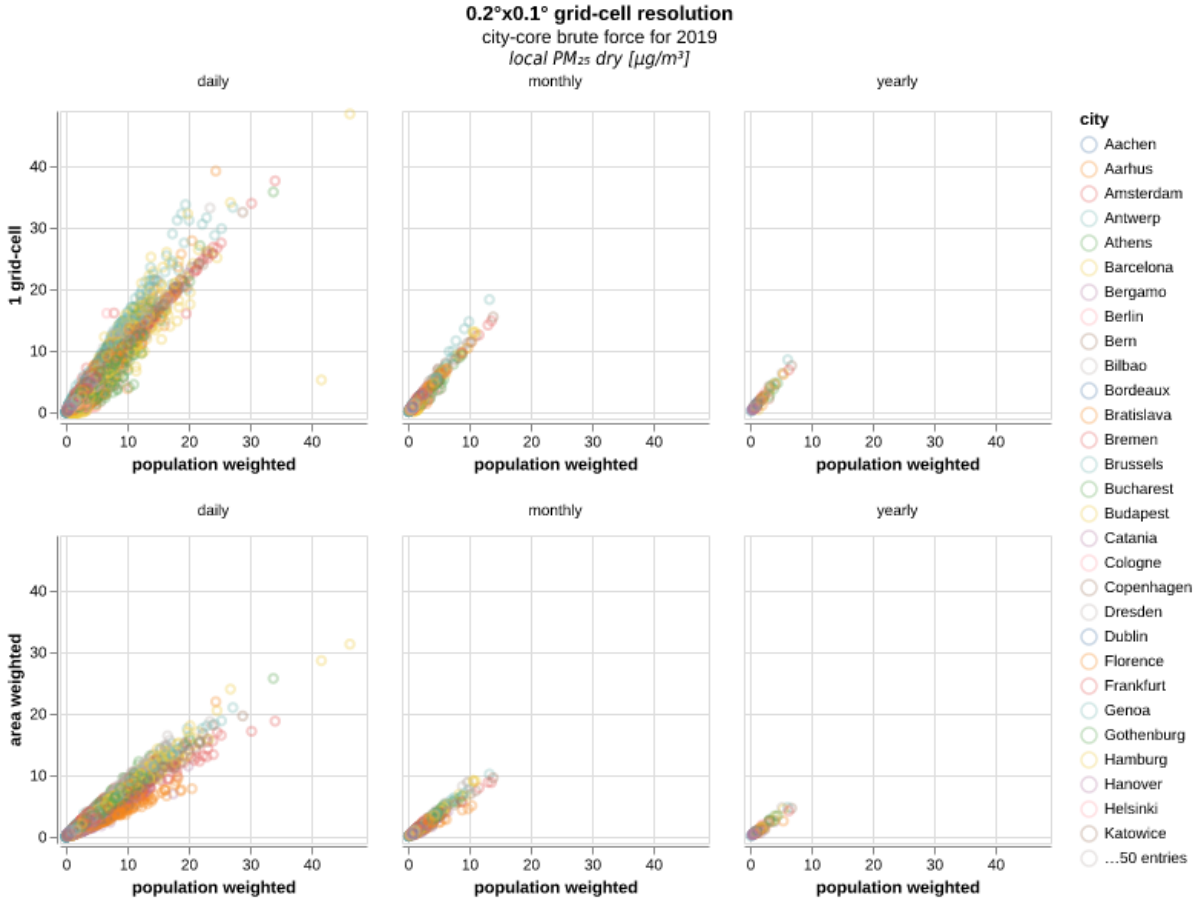


Figure 3.3: Contribution from local (core city) emissions to concentrations of $PM_{2.5}$ (ug/m3) in receptor areas, where the receptor is defined as 1 grid cell in the middle of the core city (1 grid-cell), or as area weighted concentrations in the core city (area weighted) or population weighted concentrations in core city. All the model runs have been performed in 0.2x0.1 degree resolution.

3.3.3 Differences due to resolution

When the work that is described in this deliverable started, the models for the CAMS Policy Service were run with a resolution of 0.25x0.125 degree (for the EMEP MSC-W model, for LOTOS EUROS 0.4x0.2 degree). Running the service with higher resolution costs much more CPU, and in addition it takes longer time (thus it is more difficult to comply with the restrictions on when results have to be ready on the CAMS Policy User interface every day). Therefore, we wanted to quantify how different the results would be in finer resolution to those running at the Service. We chose 0.2x0.1 degree as the finer resolution as a compromise between what is possible to achieve in practice (being able to provide results in time) and the need to improve the resolution. We compared this resolution (0.2x0.1) to 0.3x0.2, which was a compromise between the resolution run by the EMEP MSC-W model and LOTOS EUROS model.

In Figure 3.4 we compare yearly average total concentrations for $PM_{2.5}$ for all 79 cities at the two resolutions (and in 3.6 for daily, monthly and yearly resolutions). Concentrations are given both as the concentrations for the centre grid, area weighted concentrations for the core city

and population weighted concentrations for core city. Overall the results are rather similar (i.e. no systematic differences). In Figure 3.5 we show a similar plot for the local contribution to $PM_{2.5}$ in the 79 cities. Clearly, resolution is much more important for the local contribution, where the 0.2x0.1 results show systematically higher contributions for the finer resolution. This is not surprising, as the lower resolution tends to smooth out the emissions in the core city. Because local contribution is an important aspect of the CAMS Policy Support Service, we recommended to CAMS2_71 to increase resolution in the upgrade that was done in the winter 2024/2025, and the Service is now running with 0.2x0.1 degree resolution for all the products.

Figures A2 to A7 in Appendix A show results for NO_2 and O_3. For NO_2, the plots clearly show a similar picture as for $PM_{2.5}$ - that the local contribution increases with resolution. For ozone, increasing resolution increases the titration, and the negative contributions of the city becomes larger (also described in the previous section and visualized in Figure 3.1). Figure A4 in the Appendix shows that for the cities with the most titration (and lowest yearly annual ozone concentration), for instance Rotterdam, Amsterdam and Antwerp, the highest resolution systematically gives lower ozone concentrations. This is the case for all cities up to yearly concentrations of about 55-60ug/m3. It might be that those effects are different in summer and winter (e.g. titration is largest in winter), but we have not looked further into such effects here.

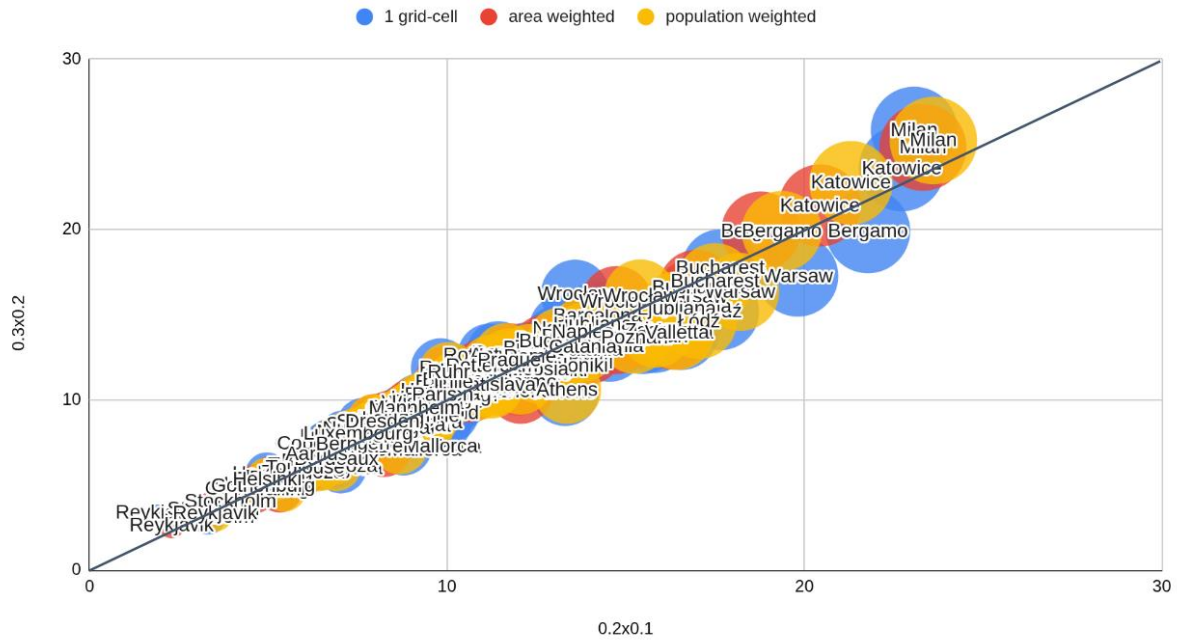


Fig 3.4: Total contribution/concentration (ug/m3) to $PM_{2.5}$ concentrations in European cities (x-axis for 0.2x0.1 degree resolution, y-axis for 0.3x0.2 degree resolution). The city source areas are the core city area, but the receptor area is either the mid-grid cell, area weighted concentrations for the core city or population weighted concentrations for the core city. The size of the circles represent the total concentrations.

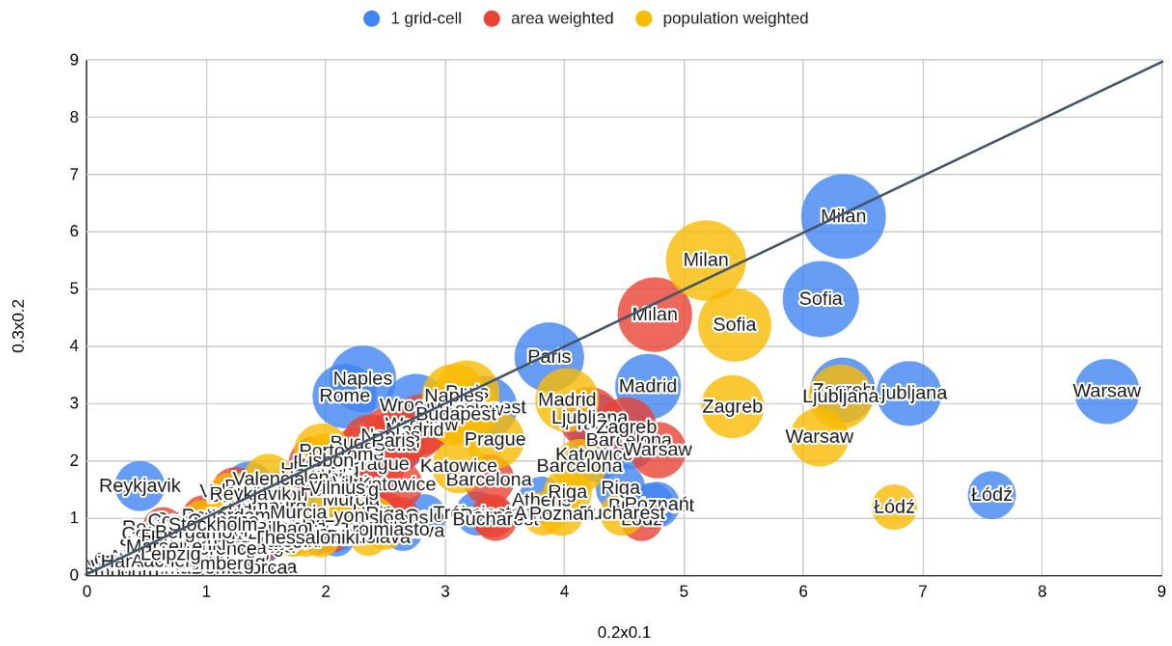


Figure 3.5: Local contribution (ug/m3) to $PM_{2.5}$ concentrations in European cities (x-axis for 0.2x0.1 degree resolution, y-axis for 0.3x0.2 degree resolution). The city source areas are the core city area, but the receptor area is either the mid-grid cell, area weighted concentrations for the core city or population weighted concentrations for the core city. The size of the circles represent the size of the local contribution.

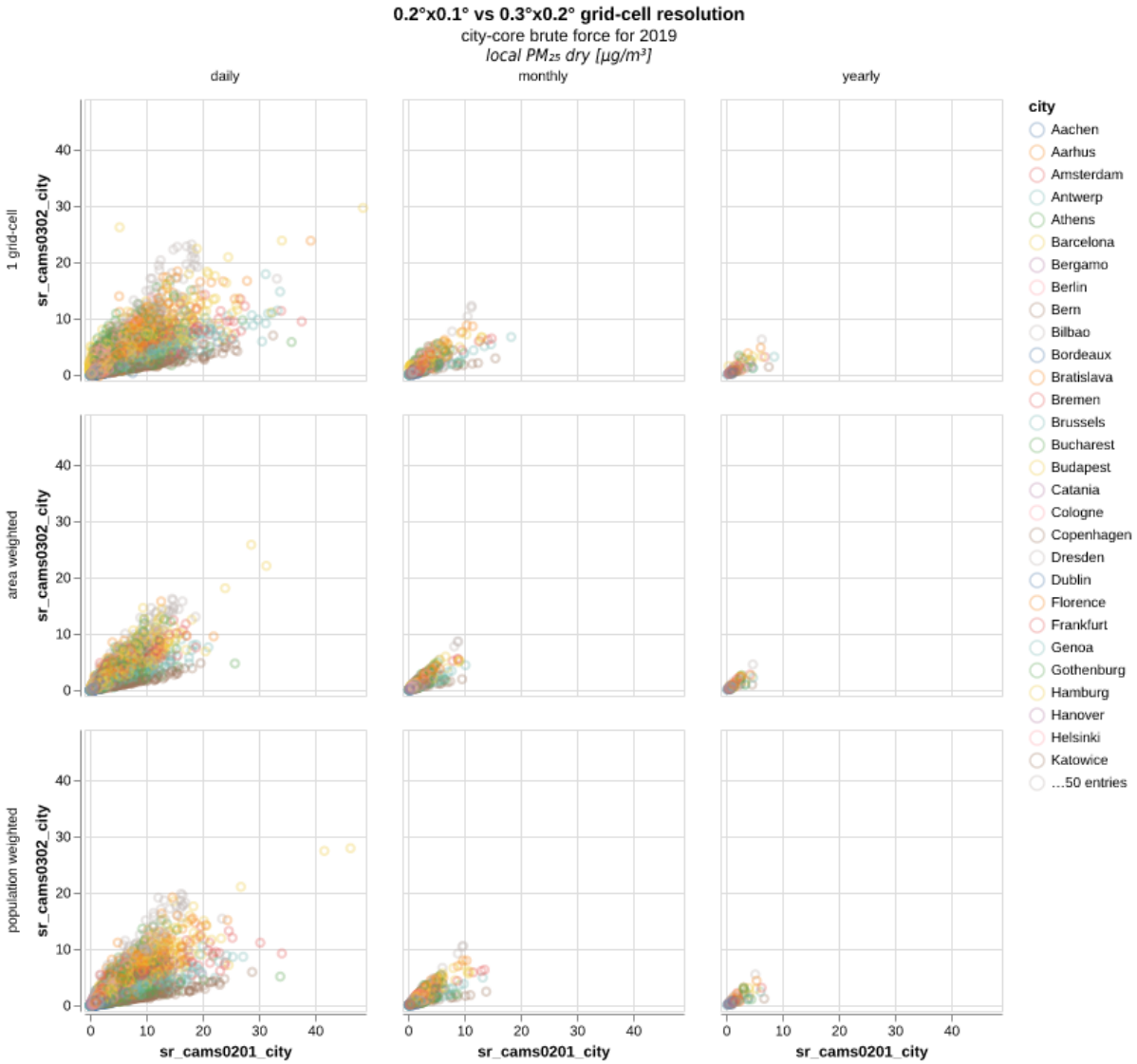


Figure 3.6: Effect of resolution on local $PM_{2.5}$ contributions (ug/m3). The source area is always the core city area, the receptor area is either the centre grid cell, area weighted concentrations in core city or population weighted areas in core city. Left column: Daily data for 2019, middle column: monthly data for 2019, Right column: Yearly average data.

3.3.4 Conclusions

We have investigated how sensitive the local contribution to the city is (i.e. how much of the pollution in the city that the city is responsible for itself) to the definition/choice of source and receptor area for the cities run in the CAMS Policy Support Service. We have also analyzed the sensitivity of the results to the resolution of the chemistry-transport model. We found that:

- The calculated local contribution is very sensitive to both resolution and geographical definition of source & receptor region, as well as to the averaging method employed within the receptor area
- Area weighted local contributions (as in the current CAMS Policy Support Service)

are systematically lower than population weighted local contributions (as in GAINS) which again are systematically lower than using the centre grid cell as the receptor (similar to SHERPA & JRC Urban $PM_{2.5}$ atlas). Therefore, the current CAMS Policy Support Service systematically attributes a smaller responsibility to the cities for its own pollution - for $PM_{2.5}$, PM_{10} and NO_2 than other policy relevant products. For O_3, the titration is highest when using the centre grid receptor and lowest for area weighted receptors.

- Higher model resolution increases the modelled local contribution systematically, by a factor 2-3 for $PM_{2.5}$, PM_{10} and NO_2 (when increasing resolution from 0.3x0.2 degree to 0.2x0.1 degree). For ozone, the titration by local city emissions is halved. However, the total concentrations are not consistently different due to resolution, except for ozone, where annual ozone is systematically lower in cities with high titration.

From a policy perspective it is not clear which definitions of source and receptor that should be used. On one hand one might argue that using the centre grid (or the grid with maximum concentration) as receptor is most relevant, as this would ensure that the sources that are most important for the worst areas are targeted when planning mitigation actions - in line with achieving the limit values set in the Ambient Air Quality Directive. On the other hand, using the population weighted concentrations would be most relevant when policy is acting on reducing health effects overall. For the source area, it seems most reasonable to choose the core city area (the administrative area), as this is the area that city authorities can act on. On the other hand, in most cities there are a lot of emissions in the surrounding areas (FUA) that are contributing to pollution in the city, and knowledge about the size of this contribution would be useful. In the end the decisions should be made by the users, and the results discussed here are planned to be presented and discussed at several meetings and workshops during fall and winter, e.g. FAIRMODE and the next CAMS Policy User workshop.

3.3.5 References

Amann, M., Bertok, I., Borken-Kleefeld, J., Cofala, J., Heyes, C., Höglund-Isaksson, L., Klimont, Z., Nguyen, B., Posch, M., Rafaj, P., Sandler, R., Schoepp, W., & Wagner, F. (2011). *Cost-effective control of air quality and greenhouse gases in Europe: Modeling and policy applications*. Environmental Modelling & Software, 26(12), 1489–1501. https://doi.org/10.1016/j.envsoft.2011.07.012

Urban Audit, 2021. https://ec.europa.eu/eurostat/web/gisco/geodata/statistical-units/urban-audit, European Commission – Eurostat. (2021). Urban Audit: Dataset on Demographics and Housing in European Cities. Retrieved from https://ec.europa.eu/eurostat.

CAO4, 2025. European Commission: Directorate-General for Environment, e-misia, EMRC, IIASA, Logika Group, MET Norway, RIVM and TNO, *Support to the development of the fourth clean air outlook – Final report*, Publications Office of the European Union, 2025, **https://data.europa.eu/doi/10.2779/8768689**

Florczyk, Aneta; Corbane, Christina; Schiavina, Marcello; Pesaresi, Martino; Maffenini, Luca; Melchiorri, Michele; Politis, Panagiotis; Sabo, Filip; Freire, Sergio; Ehrlich, Daniele; Kemper,

Thomas; Tommasi, Pierpaolo; Airaghi, Donato; Zanchetta, Luigi (2019): GHS-UCDB R2019A - GHS Urban Centre Database 2015, multitemporal and multidimensional attributes. European Commission, Joint Research Centre (JRC) [Dataset] doi: 10.2905/53473144-b88c-44bc-b4a3-4583ed1f547e PID: http://data.europa.eu/89h/53473144-b88c-44bc-b4a3-4583ed1f547e

Inness, A., Ades, M., Agustí-Panareda, A., Barré, J., Benedictow, A., Blechschmidt, A.-M., Dominguez, J. J., Engelen, R., Eskes, H., Flemming, J., Huijnen, V., Jones, L., Kipling, Z., Massart, S., Parrington, M., Peuch, V.-H., Razinger, M., Remy, S., Schulz, M., and Suttie, M.: The CAMS reanalysis of atmospheric composition, Atmos. Chem. Phys., 19, 3515–3556, https://doi.org/10.5194/acp-19-3515-2019, 2019

Kaiser, J. W., Heil, A., Andreae, M. O., Benedetti, A., Chubarova, N., Jones, L., Morcrette, J.-J., Razinger, M., Schultz, M. G., Suttie, M., and van der Werf, G. R.: Biomass burning emissions estimated with a global fire assimilation system based on observed fire radiative power, Biogeosciences, 9, 527–554, https://doi.org/10.5194/bg-9-527-2012, 2012.

Pesaresi, M., Schiavina, M., Politis, P., Freire, S., Krasnodębska, K., Uhl, J. H., … Kemper, T. (2024). Advances on the Global Human Settlement Layer by joint assessment of Earth Observation and population survey data. International Journal of Digital Earth, 17(1). https://doi.org/10.1080/17538947.2024.2390454

Thunis, P., Pisoni, E., Peduzzi, E., Vignati, E., Trombetti, M. et al., Urban PM2.5 atlas – Air quality in European cities, Publications Office, 2017, https://data.europa.eu/doi/10.2760/336669

Thunis, P., Degraeuwe, B., Pisoni, E., Trombetti, M., Peduzzi, E., Belis, C., Wilson, J., Clappier, A. and Vignati, E., PM2.5 source allocation in European cities: a SHERPA modelling study, ATMOSPHERIC ENVIRONMENT, 2018, ISSN 0004-6981, 187, p. 93-106, JRC111082.

Thunis, P., Clappier, A., de Meij, A., Pisoni, E., Bessagnet, B., and Tarrason, L.: Why is the city's responsibility for its air pollution often underestimated? A focus on PM2.5, Atmos. Chem. Phys., 21, 18195–18212, https://doi.org/10.5194/acp-21-18195-2021, 2021.

4 Impact of fine scale resolution (down to urban scale) and development of local sub-grid scale correction

4.1 Introduction

In the CAMS Policy support service, the EMEP model simulates the emission and dispersion of air pollutants at a horizontal resolution of 0.2° x 0.1° (or approximately 12 km), and a vertical resolution of ca. 50 m near the surface (the lowest layer is 50 m thick). For long-transported air-pollution this resolution can be sufficient for capturing the resulting concentrations in a city. However, for locally emitted air pollution, the concentrations will vary on a much finer scale due to non-uniform emissions. Particularly for emissions close to ground level, such as from road traffic, there will be important horizontal gradients in concentrations within a city on much finer scales than the EMEP model on 0.2° x 0.1° can resolve. Moreover, the limited vertical resolution results in the model immediately mixing low-level emissions throughout the first 50 m, which can lead to underestimation of surface concentrations from sources emitted at ground level, as well as overestimation of elevated sources.

In CAMEO tasks 6.2.2-6.2.3 we investigate how the small-scale variability in concentrations within a city, and therefore also in population exposure, could be better represented in the CAMS Policy support service. We apply the uEMEP model (Denby et al., 2020) to downscale the EMEP results to a spatial resolution of 250 m using proxy data, and combine this with population density data from the Global Human Settlement layer (Pesaresi et al., 2024) to calculate the exposure distribution (at home address). As metrics for the population exposure, we will mainly focus on the population-weighted mean concentration and on the 95-% percentile (i.e. the concentration exceeded at the home address of 5 % of the city population). As definition of the city, we will use the core city polygons defined for the Urban Audit 2021 (as described in section 3.2).

Investigations are done for NO_2 and $PM_{2.5}$. The focus is primarily on NO_2, since this is the pollutant most dominated by local sources emitted near ground level, especially road traffic. Previous work also found that downscaling of NO_2 with uEMEP gave important improvements in model performance at EEA stations, while improvements in PM were limited (Mu et al., 2022).

4.2 Experimental setup

4.2.1 EMEP model

The EMEP model is run from 1. January 2019 to 28. February 2019. The setup is the same as described in section 3.2, except that the CAMS emissions have been replaced with reported EMEP emissions for the year 2019 as used in the EMEP status report 1/2023. This was chosen due to a known problem with the gridding of NOx traffic emissions in the CAMS REG AP v6.1 emission dataset that leads to too little emissions being distributed to the cities, which was found by comparing to bottom-up emission inventories (personal communication with Jeroen Kuenen). Temporal profiles of emissions are not changed.

The EMEP model also calculates and outputs local fractions, which track how much of the primary PM and NOx concentrations in the lowest model layer of each grid cell that was emitted in the cell itself and in each neighboring cell in each GNFR sector (Wind et al., 2020)[1]. This local fraction data is needed to avoid double-counting of the local contributions modelled by uEMEP.

[1]This is the simplified local fractions method, which only tracks primary emissions from neighboring grid cells. A more complex method is used in chapter 5, which includes secondary species and sensitivities to chemical reactions.

Since the EMEP model in this setup is run with zero concentration initial conditions, the first week of the simulation is not used. The period considered for analysis is therefore 8. January – 28. February 2019.

4.2.2 uEMEP model

The urban EMEP (uEMEP) model (Denby et al., 2020) is a Gaussian plume model used to downscale EMEP model output to high resolution. The $PM_{2.5}$ and NOx concentrations are split into a local contribution that is recalculated with high resolution, and a non-local (or background) concentration that is interpolated from the EMEP model. This calculation is done independently at each receptor location. So although the downscaling can be done for a grid (mapping), it can also be done for individual locations, which can be used for validation at measurement sites with a reduced computational cost. This is used to evaluate the impact of using higher resolution for dispersion of traffic emissions, as described in section 4.2.3.

The first step is to generate fine-resolution emissions of $PM_{2.5}$ and NOx. In the setup used in this study, uEMEP does this by distributing hourly gridded emission data outputted by the EMEP model to a finer resolution grid using proxies. The proxy data is open street maps for road traffic (GNFR-F), population density for residential heating (GNRF-C), and AIS data for ship emissions (GNFR-G), all available on 250 m resolution (Table 4.1). The details of these proxy data are explained in Mu et al. (2022). However, an important difference from the approach in Mu et al. (2022) is that downscaling is done for hourly emissions rather than annual total emissions. We have also used a more recent dataset for population density.

Sector	Proxy	Emission height (m)	Comment
C - Residential combustion	Population density year 2020 from the Global Human Settlement dataset (Pesaresi et al., 2024)	15 ± 10	Downloaded at 3 arcsec resolution and aggregated to a 250 m grid
F - Road transport	Open street map (OpenStreetMap contributors)	1 ± 1	Weighting of emissions using road type according to Mu et al. (2022)
G - Shipping	AIS based emission data provided by the Coastal authorities of Norway (Kystverket, 2020)	45 ± 20	Emission vertical profile based on STEAM emission model (Jalkanen et al., 2009)

Table 4.1: Proxy data used to downscale emission from each of the three sectors, and the emission height used in Gaussian plume calculations for each sector (± indicates the initial spread of the plume). Based on Table 1 in Denby et al. (2024).

The second step is to disperse these fine-scale emissions. The dispersion is modelled as individual Gaussian plumes between the receptor and each emission grid-cell within a 2x2 EMEP-cell area around it[2]. Emission height is constant for each sector and assumed at 1 m for traffic, 15 m for residential heating at 45 m for shipping, with some initial spread (Table 4.1). Receptor height is 2 m for gridded calculations and 3 m for calculation at measurement stations. For details of the Gaussian plume calculation, see section 3 in Denby et al. (2020).

The third step is to calculate the contribution from everything that is not downscaled, which is referred to as “non-local”. This includes all sources further away than the 2x2 EMEP-cell area around the receptor, in addition to all sources in GNFR sectors that are not downscaled, as well as natural sources. The non-local contribution is calculated by subtracting the local fraction data outputted by EMEP from the total EMEP concentration in the lowest model layer,

[2] This local area is always centered at the receptor. When the receptor is not at the centre of an EMEP grid cell, the 2x2 cell area will cover fractional parts of the actual EMEP grid cells.

after interpolating (area-weighted) to the receptor location. For details, see section 2.4 in Denby et al. (2020).

No chemistry is included in the local dispersion calculations, except for the NO_2-O_3 interaction, which is modelled using the weighted timescale approach documented in section 3.4 of Denby et al. (2020). This means that the uEMEP model will modify the contribution from primary PM, while the contribution from secondary PM remains unchanged and is included in the "nonlocal".

4.2.3 Calculation areas

uEMEP is run for the core city of each of the 79 cities shown on the CAMS Policy support service. For studying population exposure statistics in a city, downscaling is done for a receptor grid of 250 m resolution. The projection used for the grid is the European ETRS89-LAEA projection (EPSG: 3035), commonly used for European mapping. For each city, a rectangular area covering the full core city polygon is downscaled, but when calculating statistics, only the 250-m cells whose centre is inside the polygon are included. Population-weighted mean and percentiles of population exposure are calculated by weighting the concentration in each 250m x 250m grid-cell within the core city polygon with the GHS population density data. This high-resolution population sampling is also done when statistics of the original EMEP concentration are calculated (i.e. EMEP grids partly within a core city polygon will be weighted with the population of that cell that is within the polygon).

For model evaluation against EEA stations (see section 4.5), downscaling is calculated for the locations of each station. For these station calculations, we also perform separate uEMEP calculations where the Open street map is sampled at 50 m resolution rather than 250 m. This strongly increases computational cost, but may improve the representation of traffic stations in particular. The residential heating and shipping are modelled at 250 m resolution in both cases. The validation against low-cost $PM_{2.5}$ sensors (section 4.6) is instead done by interpolating uEMEP results from the mapping runs at 250 m resolution, since these stations are more numerous and it was decided to include them in this study at a late stage of the work.

4.3 Average effect of downscaling

4.3.1 NO_2

Figure 4.1a shows the population-weighted mean NO_2 concentration averaged over the 52-day period in each core city. It also shows how much of the NO_2 concentration is attributed to the local downscaled emissions from each of the three downscaled sectors and to the nonlocal contribution[3]. In most cities, the downscaled emissions contribute more than half the NO_2 concentration. Road traffic is the most important of these in most cities, though residential combustion also gives an important contribution. Shipping is important only in a smaller number of cities.

Figure 4.1b shows the relative increase in average NO_2 concentration due to downscaling. Averaged over all cities, the increase is 40 %, but it varies importantly between cities, with some cities having close to no increase at all and others more than doubling. Figure 4.1b also shows the relative increase in NOx, which is significantly higher than the increase in NO_2 for some cities. This is due to the nonlinearity of NO_2 with more NOx already present, less of the emitted NO will react with ozone to NO_2. The differences in increase in NO_2 and NOx can be

[3] The NO2 source contributions are calculated from the NOx source contributions in the following way: The nonlocal NO2 contribution is calculated by photostationary approximation after removing the local NOx contribution from the downscaled sectors. Then the local NO2 contributions are calculated by repeating chemistry calculations for when one sector is removed at a time, finally normalising the local contributions so the sum of local and nonlocal NO2 contributions add to the total downscaled NO2.

much larger than shown in the figure when considering areas and time periods with particularly high NOx concentrations.

In Figure 4.1c-e we investigate the increase of each of the downscaled sectors individually. For this we consider NOx rather than NO_2. This is because the sector-wise increase in NO_2 is not trivial to define, due to interaction between the sectors from nonlinear ozone chemistry, while NOx can be easily split into sectors since NOx is linear in uEMEP. We compare the downscaled contributions calculated by uEMEP to the original local EMEP contributions. These EMEP contributions are what uEMEP replaces, and they are calculated by spatial interpolation of the local fraction data. Averaged over all cities, the traffic contribution is increased by 154 %, the residential heating by 96 %, while shipping is reduced by 59 %. For shipping, we have excluded cities where the shipping contribution in EMEP is less than 0.5 µg/m3, to avoid these insignificant cases to influence the city mean. The combined effect of downscaling the three sectors is an increase of 54 % in total NOx concentration, which becomes a 40 % increase in total NO_2 concentration. Although the general pattern of an increase in traffic and residential heating contribution and a decrease in shipping contribution is consistent among the 79 cities, there is a lot of variation in the magnitude of the increase.

These differences between cities in the effect of downscaling can have several reasons. First, the correlation between population density and downscaled concentrations could vary between cities, especially for traffic whose proxy is not population density. If we redo the comparison for a simple area-mean over the city instead of population-weighted mean, the average increase from downscaling is smaller (traffic +113 %, heating +32 %, shipping -67 -%, total NO_2 +27 %), but most of the differences between cities remain (not shown).

Second, the different cities could have different meteorological conditions affecting the dispersion of pollution, in particular the vertical stability, whose effect is better captured in uEMEP than in EMEP due to the vertical resolution (see examples in section 4.4).

Third, the subsampling of EMEP grid cells partly within the city polygon can itself lead to higher average concentrations in the cell after downscaling, since in such grid cells the emissions can be expected to be mostly in the fraction of the cell that is inside the core city, leading to higher concentrations there than in the cell average. This effect might also be different between cities, depending on the size of the city and the gradient of emissions at its boundary. Many of the cities with the largest increase in local contributions are indeed small core cities, where these boundary effects might be an important factor.

4.3.2 $PM_{2.5}$

Figure 4.2 shows the same statistics of downscaling for $PM_{2.5}$ as Figure 4.1 shows for NO_2. For $PM_{2.5}$ it is simpler to distinguish the local sector contributions since nonlinear chemistry is not involved in the downscaling and local contributions here only include primary PM. Most cities either have very little local contribution or it is dominated by residential heating (Figure 4.2a). The relative increase in total $PM_{2.5}$ concentration due to downscaling is smaller than for $NO_2$for most cities, on average a 18 % increase, even though the relative increase in the traffic and residential heating sectors are on average 155 % and 114 %, respectively. Downscaling mainly has an impact on population-weighted $PM_{2.5}$ in a smaller number of cities where residential heating gives a very large contribution.

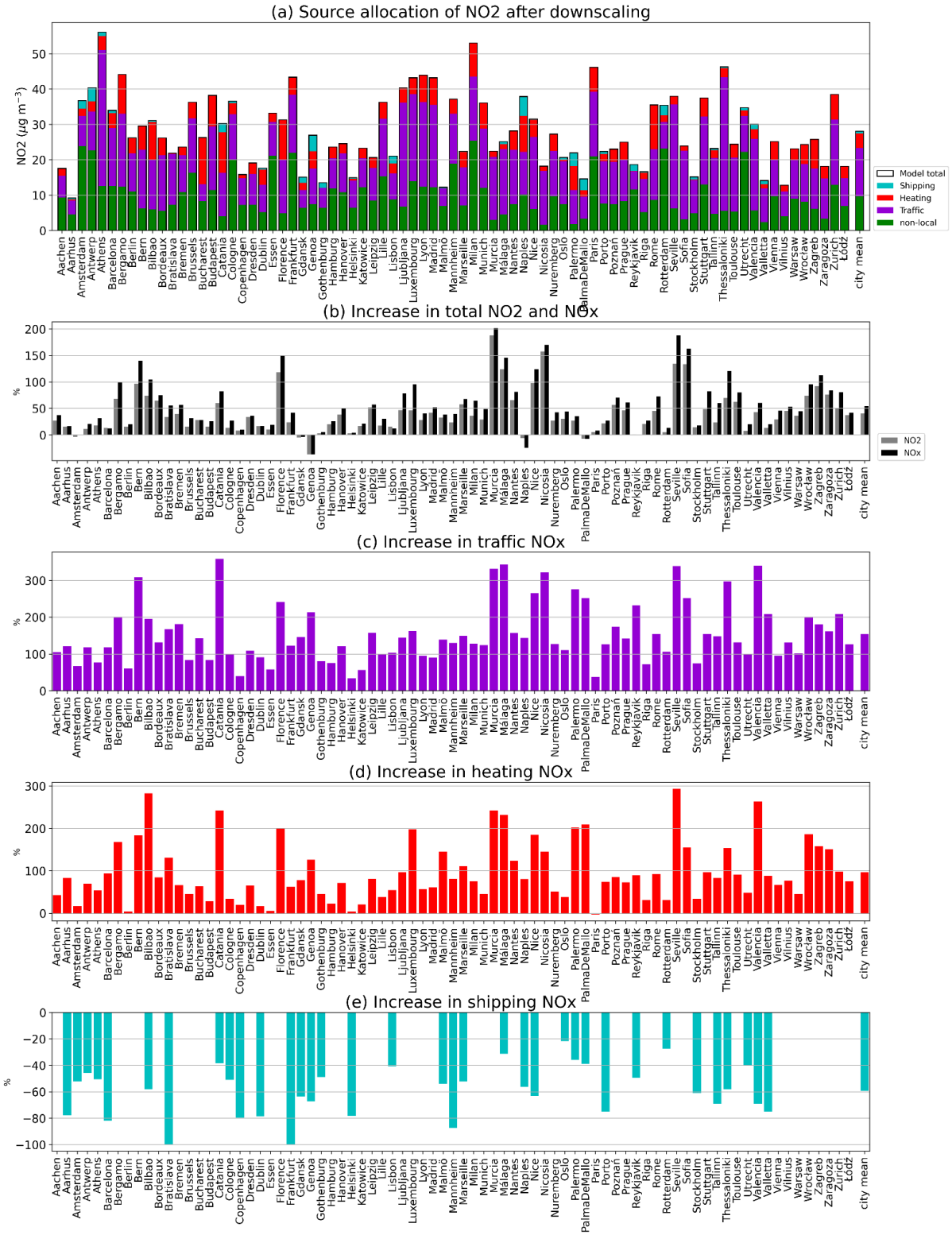


Figure 4.1: (a) Population-weighted mean NO_2 concentration in each core city after downscaling to 250 m resolution, averaged over the period 8.1.2019 – 28.2.2019, and indicating the relative importance of each downscaled local sector and the contribution that is not downscaled (nonlocal). (b) Relative increase due to downscaling in population-weighted mean of time-mean NO_2 and NOx concentration. (c)-(e): Relative increase in downscaled local sector contributions to NOx (population-weighted mean of time-mean), relative to spatially interpolated EMEP local fractions. For shipping, only cities with at least 0.5 μg/m3 original contribution are shown.

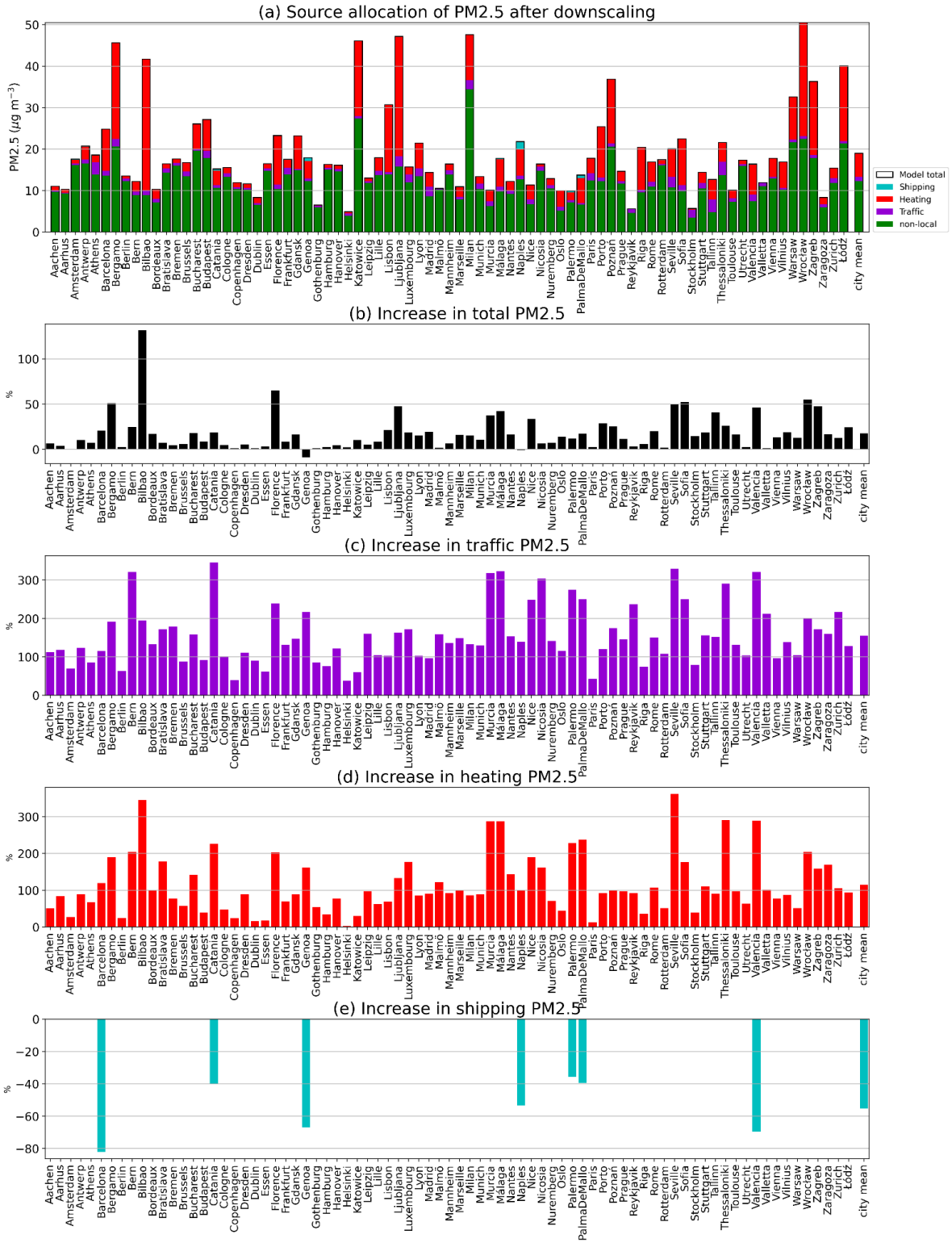


Figure 4.2: (a) Population-weighted mean $PM_{2.5}$ concentration in each core city after downscaling to 250 m resolution, averaged over the period 8.1.2019 – 28.2.2019, and indicating the relative importance of each downscaled local sector and the contribution that is not downscaled (nonlocal). (b) Relative increase due to downscaling in population-weighted mean of time-mean $PM_{2.5}$ concentration. (c)-(e) Relative increase in downscaled local sector contributions to $PM_{2.5}$(population-weighted mean of time-mean), relative to spatially interpolated EMEP local fractions. For shipping, only cities with at least 0.5 μg/m3 original contribution are shown.

4.4 Time series examples for NO_2

Figure 4.3 shows an example of the downscaling of NO_2 for a 10-day period in Berlin, showing a map of time-averaged concentration before and after downscaling. Most of the core city is covered by only four EMEP model grid cells, so the EMEP model can only capture the large-scale variability, with higher concentrations in the city centre than in its peripheries. Much more details can be seen in the downscaled concentrations, particularly elevated concentrations near the main roads. While the highest concentration in the EMEP model is only around 28 µg/m3, in uEMEP some areas have a concentration as high as 45 µg/m3.

Figure 4.4a shows the time series of population-weighted mean concentration in the same period, before and after downscaling, as well as the 5-95 % range in population exposure (shaded). Observational time series from EEA stations (see section 4.5) are super-imposed, color-coded for station type. Since in particular traffic stations are located in the highest concentrations, the maximum value from uEMEP (among populated 250m x 250m cells) is also shown.

First, we can see that although the population-weighted mean concentration is higher in uEMEP than in EMEP on average, there are also many times where they are very similar to each other, or even lower in uEMEP. The 95-% percentile is also higher in uEMEP than EMEP in many cases, especially when the mean is higher, though there are also periods when it is similar to the 95-% percentile from EMEP (which normally corresponds to the EMEP grid with highest concentration). The city-max concentration from uEMEP, however, is usually much higher than uEMEP 95-% percentile. This can be because the areas of strongly elevated NO_2 concentration along the main roads, though very clearly visible in Figure 4.3b, may not constitute as much as 5 % of the population.

Even with the downscaling, it seems the model underestimates the NO_2 concentration in this particular episode. This is also found from direct comparison to observations at station locations (not shown). This is likely not only due to local sources, though, since there is also underestimation at night, indicating that background concentrations are also too low.

Figure 4.4b further investigates the population-weighted mean NO_2 concentration and its increase by downscaling, and splits the downscaled concentration into contributions from downscaled sources and the nonlocal (similar to Figure 4.1a). As noted above, the concentration is strongly increased by downscaling in some periods, while in others it is close to what we get directly from the EMEP model. The increase from downscaling seems to occur especially during the afternoon, when also the contribution from local road traffic peaks. There is also a lot of variability from day to day, with some days having very little effect of downscaling and other days having a large increase. Residential heating contribution also has a peak slightly later in the evening, which might also contribute to the larger effect of downscaling.

One possible reason for the strong time-variability in the effect of downscaling is the stability. When the boundary layer is stable, pollution emitted near the ground will mix vertically more slowly than when it is unstable. While EMEP will immediately mix the emissions throughout the lowest layer (~50 m thick), uEMEP accounts for the mixing taking longer when it is stable through the Gaussian plume formula, which would give rise to higher surface concentrations in uEMEP than EMEP in stable conditions. To explore the stability as an explanation, Figure 4.4c shows the uEMEP-to-EMEP ratio in population-weighted mean contribution to NOx from road traffic (i.e. a value of 1 means they are the same), and super-imposed is the stability function for heat as outputted by EMEP. The stability function is bigger than 1 in stable conditions and smaller than 1 in unstable conditions. We see there is a clear relationship between the two, with the large increases in NOx traffic contribution corresponding to the most stable conditions. The sharp increases in the ratio in the afternoon coincides with the boundary layer transitioning from stable to unstable. In the first few days of the period, the stability was not so strong, and then downscaling only weakly increased the NOx traffic contribution, sometimes even decreasing it. The squared correlation coefficient between the stability function and the NOx traffic relative increase in the time series shown has a coefficient of

determination of r^2=0.60. However, the relationship is not perfect. For example, there is a tendency for stability to remain high late in the evening, while the NOx traffic increase goes down. This is also when the traffic emissions drop, so one possible explanation is that local NOx emitted earlier in the evening is still present in EMEP; this contribution from earlier cannot be included in uEMEP since the Gaussian plumes are in steady-state with no memory of emissions at previous times.

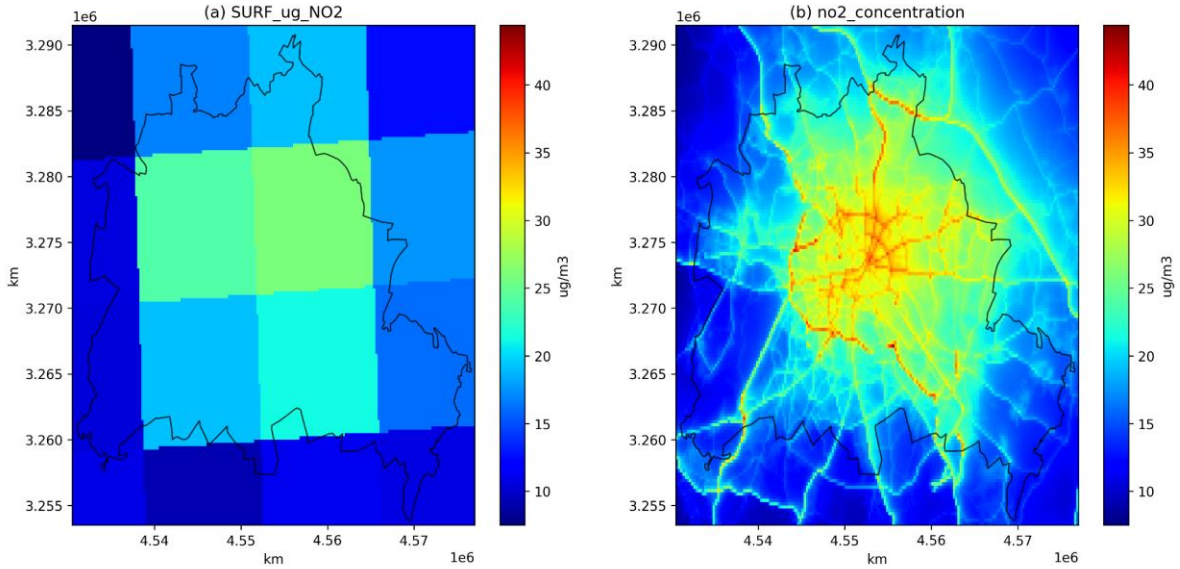


Figure 4.3: NO2 concentration averaged over the period 10.–19. February 2019 in Berlin: (a) Original concentration in EMEP with no downscaling. (b) Downscaled concentrations from uEMEP at 250 m resolution. The core city boundaries are indicated by the black polygon.

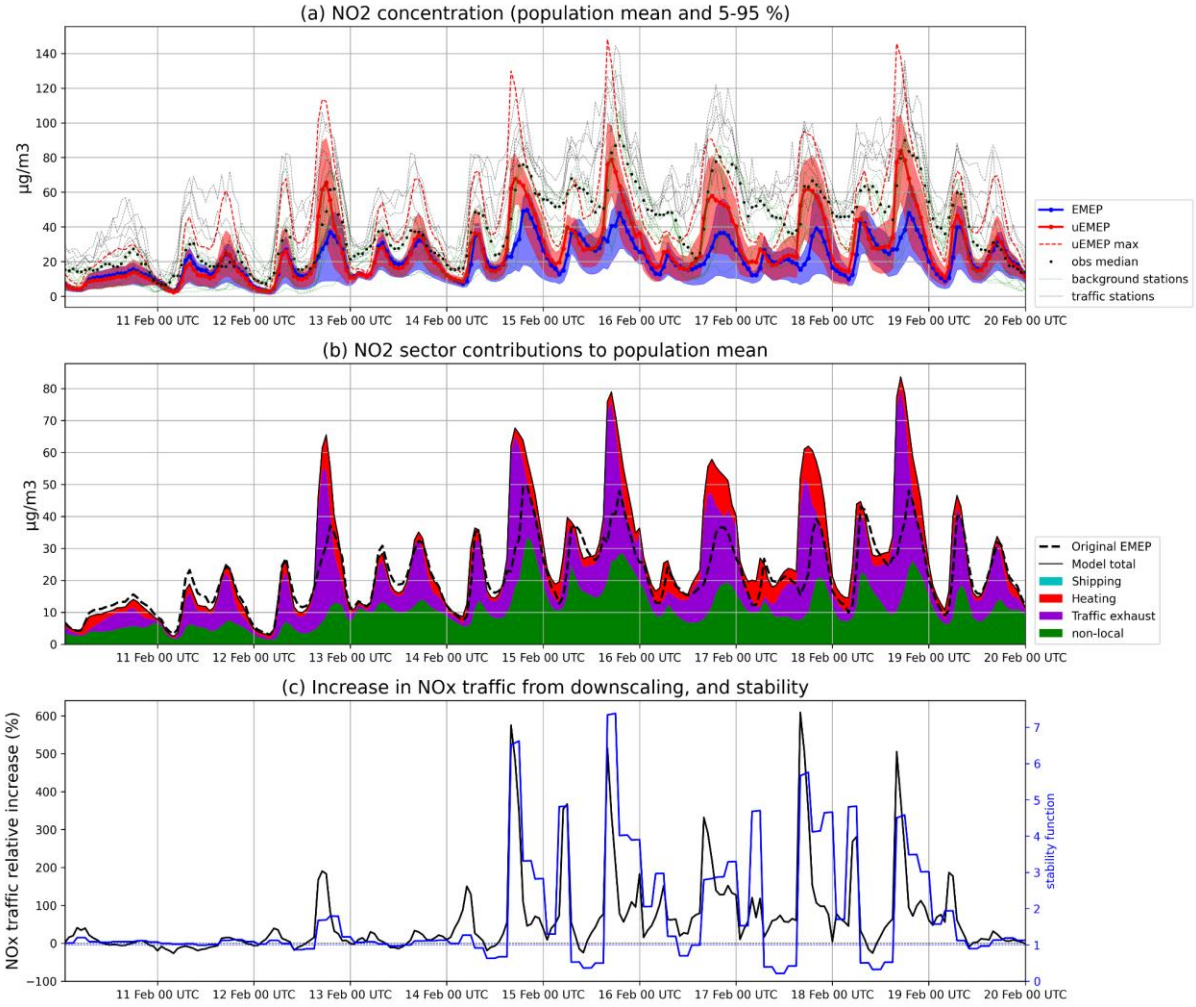


Figure 4.4: Modelled NO_2 concentration in Berlin core city hour-by-hour 10.–19. Feb 2019: (a) Population-weighted mean (thick line) and 5-95% percentile range of population exposure (color segments) in EMEP (blue) and uEMEP (red), as well as the uEMEP populated receptor grid-cell with the highest concentration (dashed red line). Super-imposed are measurement stations (thin dashed lines) and median of stations (dots). (b) Population-weighted mean concentration in uEMEP, split into contributions from each downscaled sector and from nonlocal, as well as population-weighted mean concentration in EMEP (dashed line). (c) The ratio of EMEP and uEMEP NOx contribution (population-mean) from local road traffic (black) super-imposed on the time series of the stability function at 10 m calculated in the EMEP model (blue).

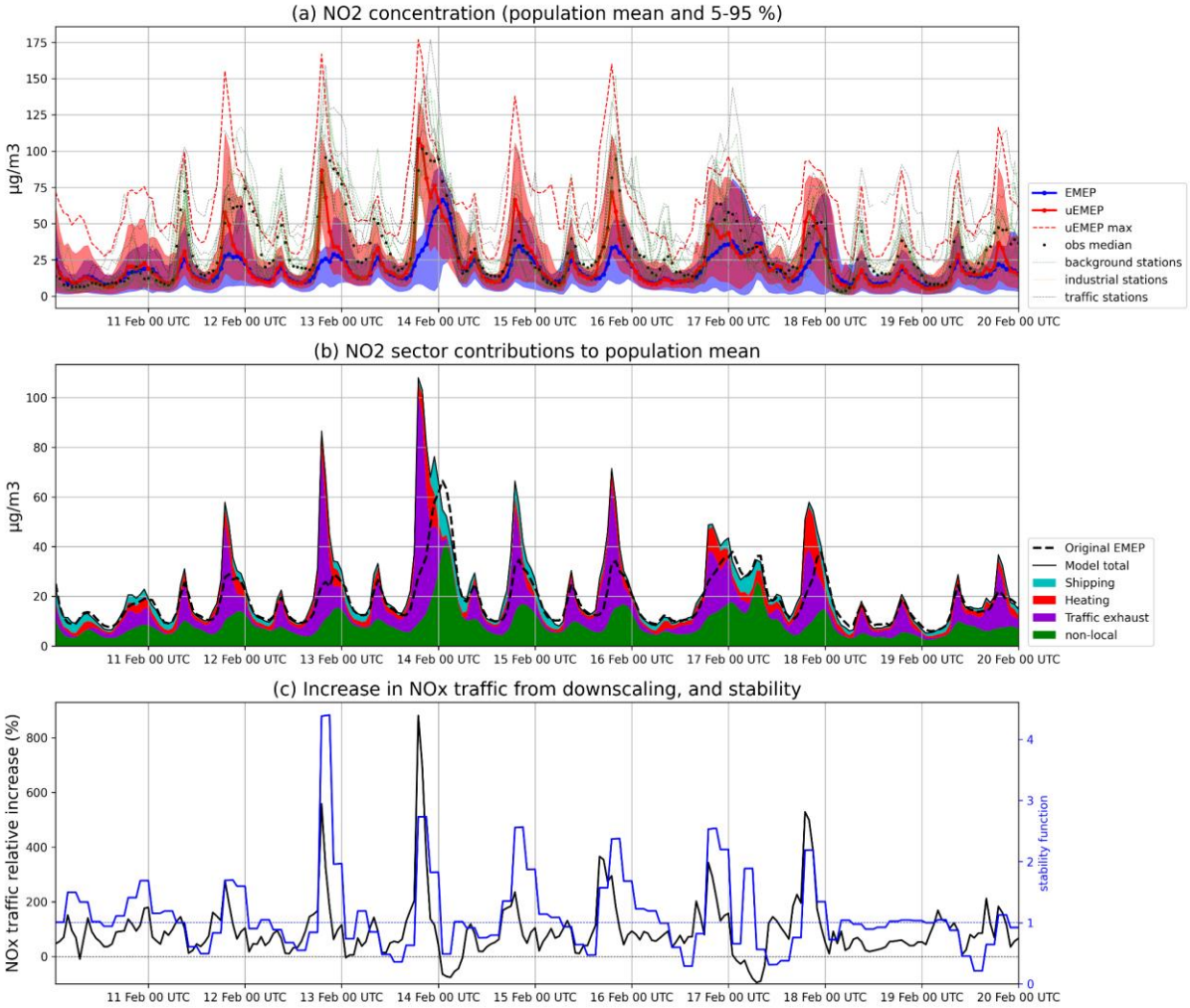


Figure 4.5: Same as Figure 4.4, but for Lisbon core city.

Figure 4.5 shows the same time period for Lisbon. Here we also see that the model tends to be lower than the observed NO_2 concentrations, but after downscaling, the population-weighted mean concentration becomes significantly higher in several of the afternoon peaks that can be seen in the station observation median. The uEMEP max concentration also peaks relatively close to the station with highest observation (although it could be different locations inside the core city). Also for Lisbon, traffic is the most important of the three downscaled sectors, although residential heating and shipping sometimes contributes significantly, too. Just as in Berlin, the increase in NOx traffic contribution from downscaling varies importantly in time and peaks in the afternoon at the same time as the strong increase in stability (Figure 4.5c). In this case r^2=0.34, so not as strong a relationship as in Berlin.

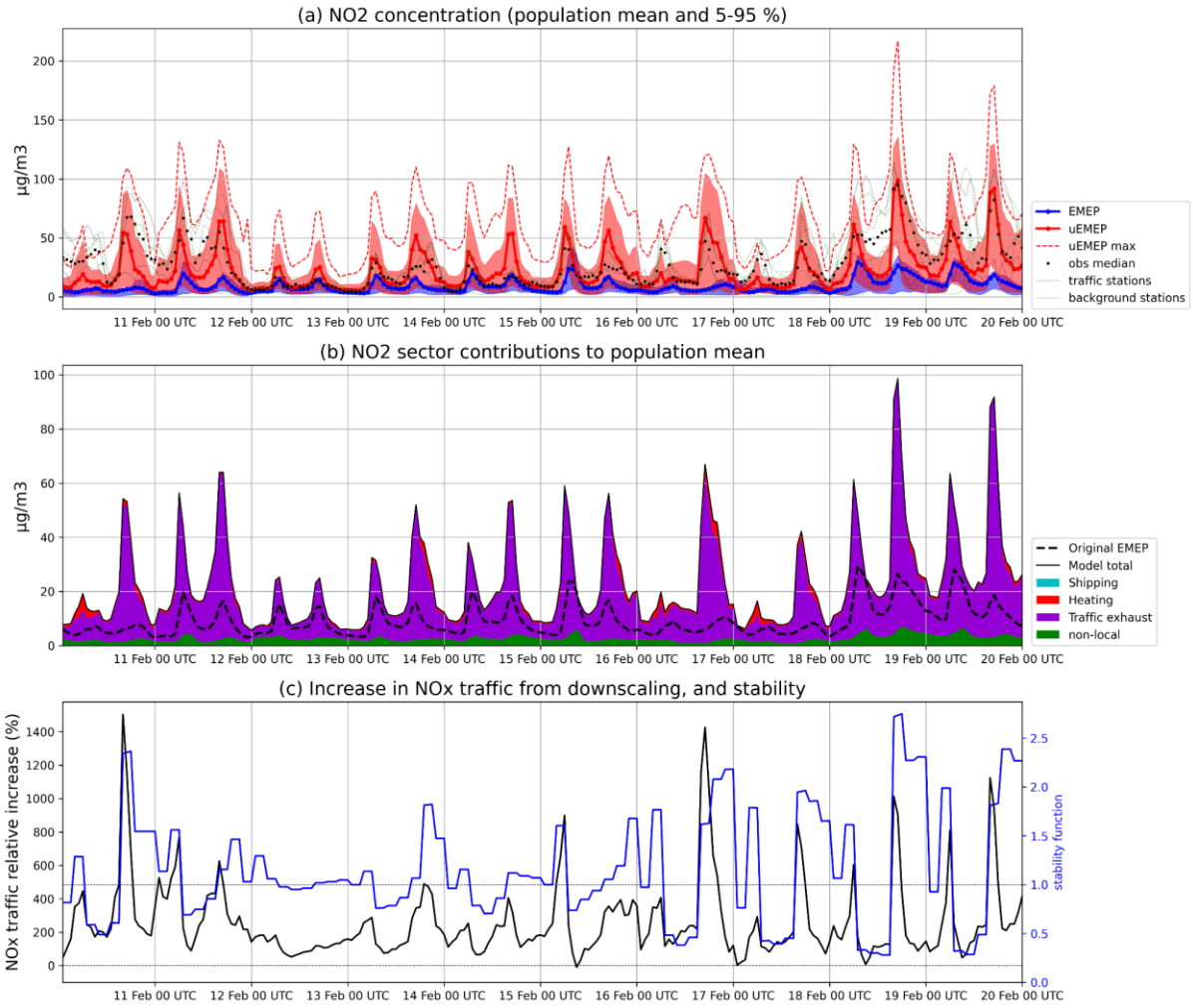


Figure 4.6: Same as Figure 4.4, but for Sofia core city.

A third example is shown in Figure 4.6, for the core city of Sofia, which is one of the cities with the largest increase in NO_2 from downscaling. In this case, the uEMEP concentration is higher than the EMEP concentration for the whole period, but the increase is very variable in time also here. The peaks in uEMEP concentration appear to correspond well with the peaks in median observed concentration (Figure 4.6a). The variability in the increase in traffic NOx contribution also has an association with stability in this case (Figure 4.6c).

Looking at all the 79 cities, we can see a similar association between the stability function and the variability in increase in traffic NOx in most of them (not shown). However, it is clear that stability alone cannot reproduce the time variation in the relative increase in NOx traffic contribution, and the relationship with stability is more of qualitative nature than a simple one-to-one dependency. This is not surprising, since other factors such as boundary layer height, wind speed and wind direction will also affect the downscaling. Due to the important time-variation in the effect of downscaling, it seems overly simplistic to apply a constant correction factor for each city at all times in the CAMS Policy support service.

4.5 Model evaluation against EEA stations

The concentrations are evaluated against measurement data from the validated and reported dataset for 2019, received from the EEA in early 2024. All stations located within the 79 core cities are included for validation if they have a data coverage of at least 75 % in the period 8.1.2019 – 28.2.2019. This gives a total of 491 stations for NO_2 and 163 for $PM_{2.5}$. However, many cities have a small number of stations. For a more robust validation of city spread, only cities with at least 5 stations, including at least 1 traffic station and 1 background station, are

included. This leaves 30 cities for NO_2 and 12 for $PM_{2.5}$. For each hour with measurements, the measured concentration is compared to the corresponding modelled concentration at the station, both before downscaling (surface concentration output from EMEP) and after downscaling (uEMEP).

4.5.1 NO_2

Figure 4.7 shows the validation of NO_2. Figure 4.7b shows that uEMEP gives an increase in the mean NO_2 concentration in most cities, which in most cases brings the concentration closer to the observed mean, although the NO_2 concentration is still underestimated in average after downscaling. The average bias goes from -12.0 µg/m3 in EMEP to -5.3 (-4.4) µg/m3 in uEMEP run at 250 m (50 m).

Figure 4.7c shows the variability between stations within each city, quantified as the difference between the station with highest and lowest period-mean concentration. In mean over the cities, the observed spread is 32.8 µg/m3, which is almost as high as the observed average (38.3 µg/m3), so variability within the cities is clearly important. In most cities, the observed variability is larger than in the model, even after downscaling. However, the spatial correlation between the stations (Figure 4.7d) is strongly improved by the downscaling for nearly all the cities, and on average it is improved from 0.26 to 0.64 (0.71) when using 250 m (50 m) resolution. It is as expected that the 50 m resolution improves the representation of traffic stations relative to 250 m resolution. Note that the clear outlier Zargoza, where spatial correlation is negative, has rather small spatial spread as well.

Figure 4.7e shows the station mean of temporal correlation in each city, which is an indication of whether the higher and lower concentrations occur at the right times at each station. The temporal correlation is not as strongly affected by the downscaling as the spatial correlation. It slightly improves on average, from 0.60 in EMEP to 0.66 in uEMEP, although in a few cities the temporal correlation gets worse after downscaling.

Finally, Figure 4.7f shows the station mean of temporal root-mean-square error (RMSE) averaged over all stations, which should give an overall indication of how well the model is able to represent the measurement time series at the stations. In most cities, the RMSE is lower (and therefore better) in uEMEP compared to EMEP, and in many cases even a ~20 % reduction. Averaged over the cities, the RMSE drops from 18.7 µg/m3 in EMEP to 16.0 (15.7) µg/m3 in uEMEP at 250 m (50 m) resolution.

For air quality it is in particular the periods with highest concentrations that are important to model correctly. Therefore, the validation is repeated after selecting only the 20 % of the time period that has the highest NO_2 concentrations. This selection is done using the median of observations from all stations in each city every hour and ranking the hours by this observed median. Thus, the data selection is different for each city and may contain isolated hours from different days. 75 % data coverage is checked after the selection of the period with 20 % highest concentrations, which means the set of stations included for validation could be slightly different from the validation of the full dataset.

Figure 4.8 shows the resulting validation for this data subset. The observed mean is much higher in this subset, on average at 68.3 µg/m3 compared to 38.3 µg/m3 for the full period. The average bias is -28.0 µg/m3 in EMEP and -11.5 (-10.2) µg/m3 in uEMEP at 250 m (50 m) resolution (Figure 4.8b), which is a similar relative reduction in the bias as for the full dataset (Figure 4.7b). Spatial correlation (Figure 4.8d) is similar for this data subset as it was for the full dataset, with important improvements from downscaling. Temporal correlation (Figure 4.8e) is much worse for the subset of highest concentrations compared to the full subset (both in EMEP and uEMEP), dropping from around 0.6 to 0.3 on average. This is probably because the variability is less dominated by the diurnal cycle of traffic emissions when only the 20 % hours with highest concentration are considered, and we can expect the diurnal cycle to be better captured in the model than the variability between peak concentrations at different days. We still see that the downscaling improves the temporal correlation on average, though again

there are several cities where it does not. But while for the full dataset, downscaling only weakly impacted temporal correlation (Figure 4.7e), the downscaling has a much stronger effect on the temporal correlation for the 20 % hours of worst concentration (Figure 4.8e). This indicates that the downscaling correctly captures an aspect of the time variability in the highest concentrations in some cities, while in others it does not. Given the strong relationship found in section 4.4 between stability and the increase from downscaling, we suspect that the reason for the differences in temporal correlation between cities could be related to whether meteorological conditions and in particular the stability are correctly represented.

Finally, the temporal RMSE is decreased in most cities also in the 20 % highest concentration subset, on average from 32.4 µg/m3 in EMEP to 25.1 (24.6) µg/m3 in uEMEP at 250 m (50 m) resolution (Figure 4.8e).

In conclusion, the downscaling of NO_2 importantly improves the bias and spatial correlation in most of the 30 cities considered for validation, both for the full dataset and for the 20 % of hours with highest concentration. Temporal correlation also improves on average, but the improvement is not as large and there are also some cities where it gets worse. Using 50 m resolution for downscaling road traffic instead of 250 m further improves both bias and spatial correlation, while it has very little effect on the temporal correlation.

4.5.2 $PM_{2.5}$

Figure 4.9 (left panels) shows the validation for $PM_{2.5}$ against EEA stations for the full period. Only 12 cities have a sufficient number of observations. The city-mean bias is improved from -2.7 to -0.9 (-0.6) µg/m3 for uEMEP with 250 m (50 m) resolution (b), but the spatial correlation is not systematically improved across the cities (d). On average, it improves from 0.05 in EMEP to 0.16 (0.19) in uEMEP using 250 m (50 m) resolution, but there are several cities where the spatial correlation gets worse after downscaling. The spatial correlation also is quite low even after downscaling and even negative in multiple cities. However, in contrast to NO_2, the observed spatial spread in the cities is only 6.1 µg/m3 on average (c), which is relatively small compared to the average concentration (18.0 µg/m3).

As for NO_2, the validation of $PM_{2.5}$ is repeated for the 20 % of the hours with the highest observed median $PM_{2.5}$ concentration in each city (Figure 4.9, right panels). For this subset, the average bias is -8.3 µg/m3 before downscaling and improves to -5.7 (-5.3) µg/m3 with downscaling to 250 m (50 m), while spatial correlation goes from 0.06 to 0.19 (0.20), but again with many cities where spatial correlation gets worse. The temporal correlation and RMSE are both getting worse with downscaling in both the full dataset and the 20 % highest concentration subperiod (Figure 4.9e-f).

Since residential heating is the most important downscaled contribution for $PM_{2.5}$, these results indicate that the population density is not a very good proxy for distributing these emissions in many of the cities. Better results might be obtained if more appropriate proxies for these emissions are found. Using annual uEMEP calculations for all of Europe, Mu et al. (2022) tested altering the proxy for residential heating in uEMEP by using a nonlinear relationship between population density and emissions or combining it with data on building density, but neither significantly improved the spatial correlation. They did, however, show that proxy data more closely linked to combustion, as available for Norway, can give significant improvements.

In addition, the performance is also limited by the original gridding of emissions. Since this study uses EMEP emissions, the emissions are mainly as reported by the countries themselves, although specifically for PM emissions from sector C (residential combustion) the emissions instead come from estimates by TNO for some countries, in the cases where the reported emissions are considered to not take into account the condensables. This is documented in the EMEP status report 1/2023, chapter 3.3. The gridding of the emissions from TNO are typically also based on proxies such as population density or wood availability, and may not account for factors such as regulation on combustion in cities. For example, PMF data in Barcelona show that contribution from biomass burning is overestimated by the models

there (Deliverable D6.3 of CAMEO). Note, however, that for our study here the Spanish emissions for GNFR C for $PM_{2.5}$ are reported by Spain and not taken from TNO.

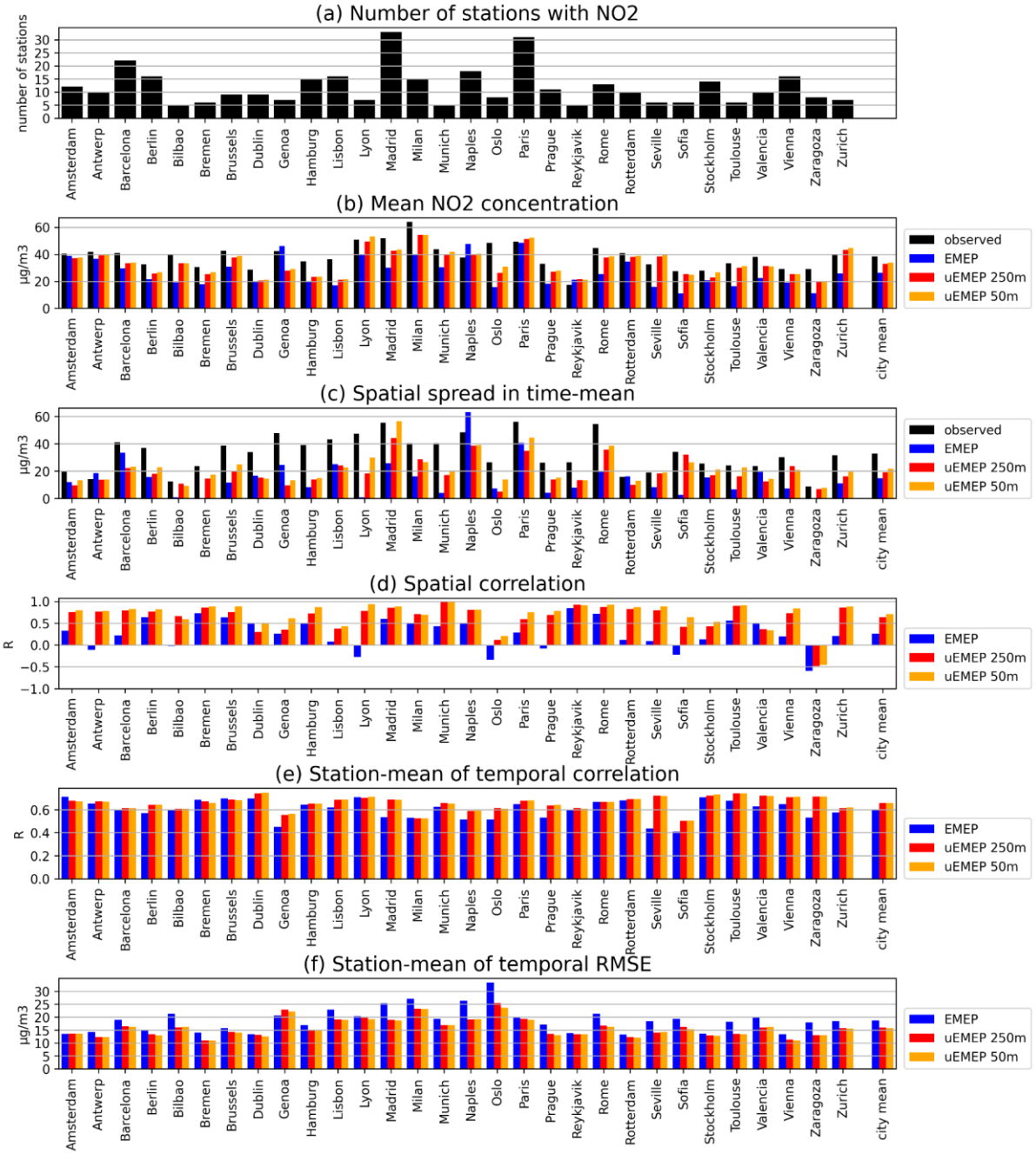


Figure 4.7: Validation of NO_2 concentration for period 8.1.2019 – 28.2.2019, for original EMEP concentrations and downscaled with uEMEP at 250 m and 50 m resolution. NO_2 is validated against EEA stations for cities with at least 5 total stations, 1 traffic station and 1 background station with >75 % data coverage: (a) The number of stations, (b) the mean concentration over the period averaged over the stations, (c) difference in mean concentration between highest and lowest station, (d) the spatial correlation (among stations in a city) of time-mean concentration, (e) the temporal correlation calculated at each station and then averaged over the stations, and (f) the temporal root-mean-square error calculated at each station and then averaged over the stations. To the far right is shown the mean over the shown cities.

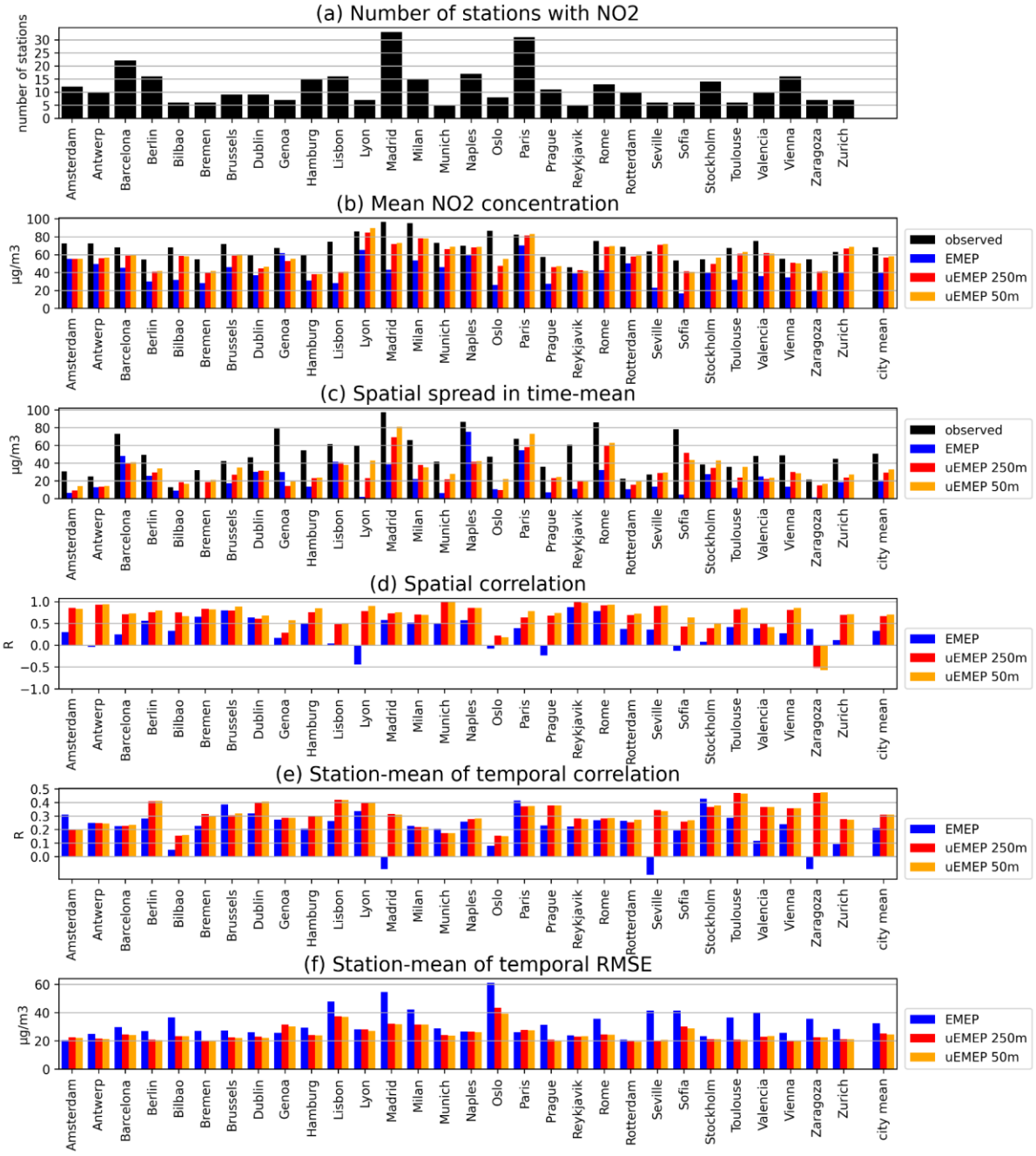


Figure 4.8: Same as Figure 4.7, but only for the 20 % of the hours in the 52-day period that have the highest NO_2 concentration (from median of the station observations).

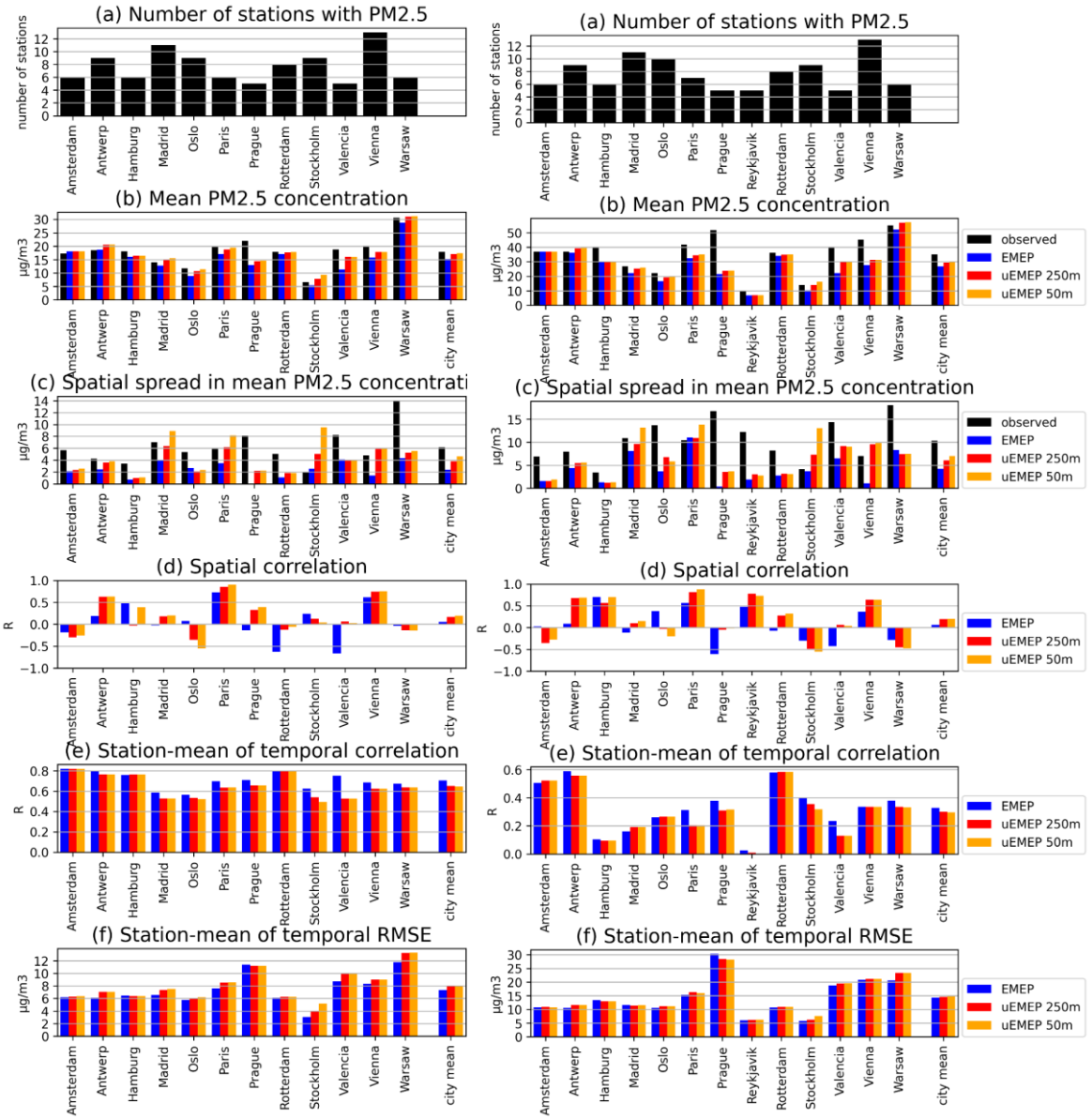


Figure 4.9: Same as Figures 4.7-4.8, but for $PM_{2.5}$, showing the full dataset to the left and the 20 % of hours with highest $PM_{2.5}$ concentration to the right.

4.6 Model evaluation against low-cost $PM_{2.5}$ sensors

The number of EEA stations available for $PM_{2.5}$ is relatively limited. As an additional way of evaluating the $PM_{2.5}$ variability within the cities, we used a recently published dataset of low-cost sensors (LCS) for $PM_{2.5}$ (Hassani et al., 2025). This dataset contains several thousand sensors from all over Europe, though mostly in Germany and surrounding countries. We have only used data with the highest level of quality flags, which means a relatively small fraction of the sensors have 75 % data coverage for the 52-day period. Only these stations have been considered. Just as for the validation against EEA stations, we also consider the hours with 20 % highest concentrations (from median of LCS available at each hour) in each city as a subset. This results in 15 cities with at least 10 sensors with 75 % data coverage for the whole dataset, and 13 cities for the subset of 20 % highest concentrations. For these cities, the gridded uEMEP concentrations at 250 m have been interpolated to the sensor locations and compared to the LCS data. Results are shown in Figure 4.10.

Similar to what we found in the EEA station validation, $PM_{2.5}$validated against the LCS data is not systematically improved by the downscaling. Spatial correlation does not systematically improve with downscaling, and the temporal correlation and RMSE both get worse on average.

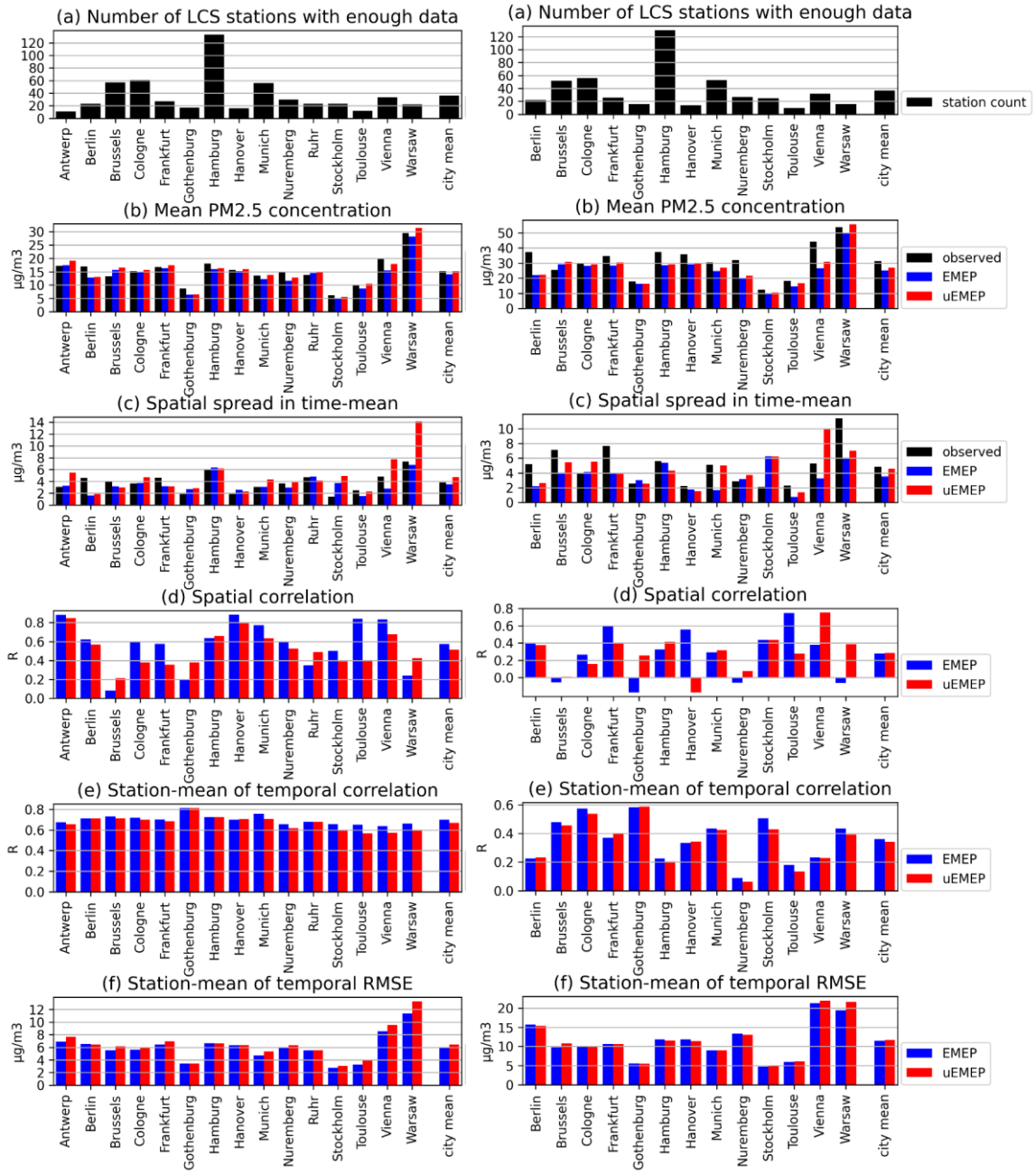


Figure 4.10: Validation of $PM_{2.5}$ concentration against Low-Cost sensors with 75 % data coverage in period 8. Jan – 28. Feb 2019 (left), and for the 20 % of this period with the highest observed $PM_{2.5}$ concentration in each city (from median of LCS observations) (right). Showing only cities with at least 10 stations with sufficient data coverage. Blue bars show original concentrations from the EMEP model, red bars after downscaling to 250 m resolution.

4.7 Downscaling effect on population exposure

Having established that NO_2 concentrations are improved from the downscaling, we will quantify how much the downscaling increases the calculated levels of population exposure. This was already shown in Figure 4.3, but only for a time-average of population-weighted mean. Figure 4.11a-b shows histograms of the occurrence of daily mean NO_2 concentrations, for the population-weighted mean and for the 95-% percentile, respectively, in all cities before

and after downscaling. There is a shift towards higher values both for the population-weighted mean and the 95%-percentile, and the distributions get wider. The limit value for daily mean NO_2 concentration defined in the new European Ambient Air Quality Directive is at 50 µg/m3, and this limit is exceeded much more frequently after downscaling, both for the population-weighted mean (238 % more days) and for the 95-% percentile (219 % more days). There are also much fewer days with concentrations below 10 µg/m3 or even 20 µg/m3.

For completeness, we also show the same histograms for $PM_{2.5}$ in Figure 10c-d. Also $PM_{2.5}$is shifted towards higher concentrations, though not as much as NO_2. The number of days above the new limit value for daily mean $PM_{2.5}$ of 25 µg/m3 increases by (only) 31 % for the city mean and by 52 % for the 95-% percentile. However, the very high values of daily mean $PM_{2.5}$ above 50 µg/m3 (and even 100 µg/m3) occur much more frequently, particularly in the 95%-percentile. Since the downscaling of $PM_{2.5}$ does not systematically improve performance, this increase in occurrence of very high concentrations might not be so reliable.

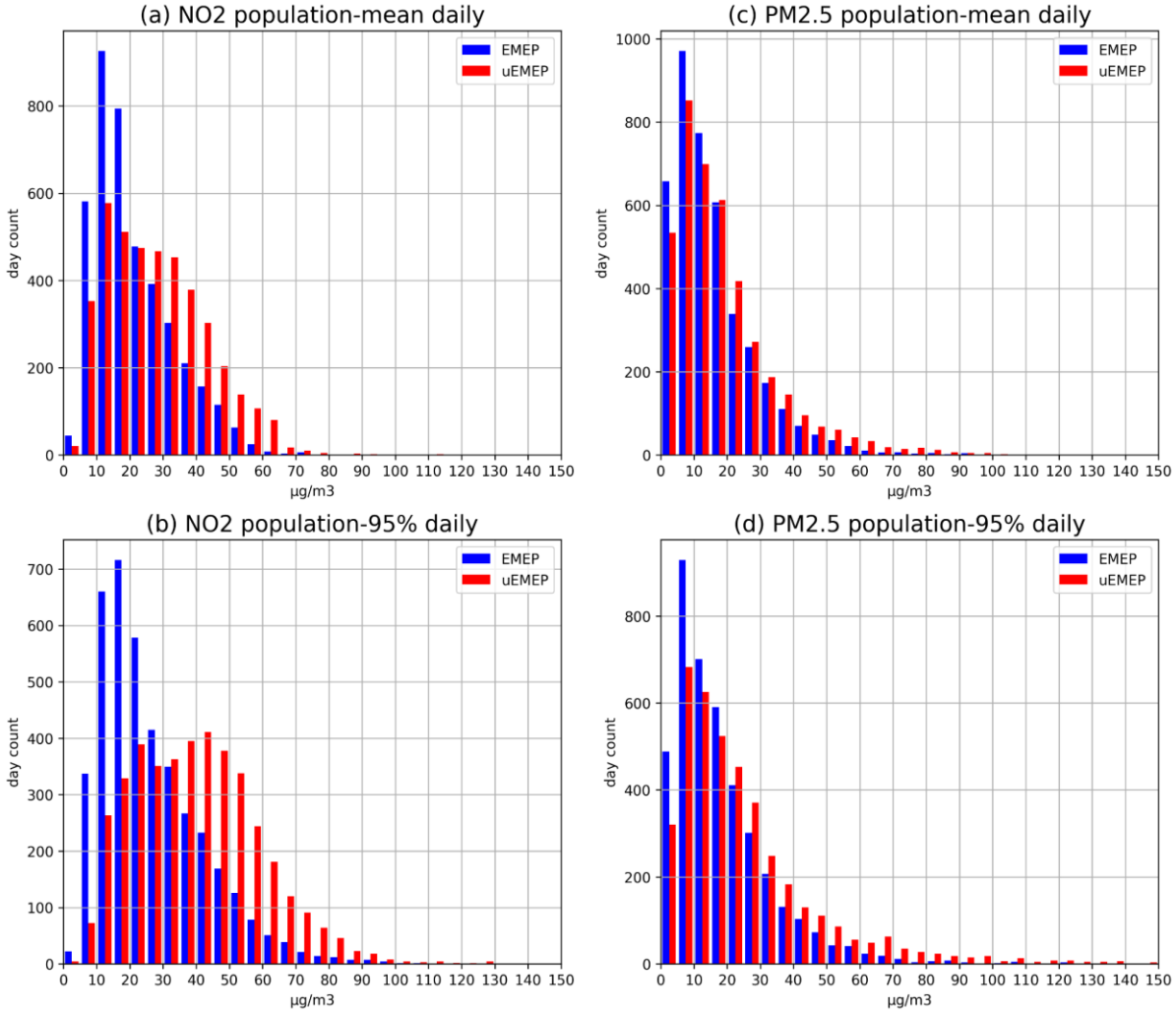


Figure 4.11: Distribution of daily mean concentrations during the 52-day period, comparing before (EMEP) and after downscaling to 250 m (uEMEP), plotted as the number of days that fall within each 5 µg/m3 bin, summed over all cities: (a) Population-weighted mean NOx concentration (b) 95%-percentile of population exposure to NOx. (c) and (d) show the same for $PM_{2.5}$. NB: There are 2 days in (b) and 8 days in (d) that exceed the 150 µg/m3 cut-off.

4.8 Recommendations for the CAMS Policy support service

We found that downscaling with uEMEP to 250 m can give important additional information on exposure to the local NO_2 from road traffic, both by correcting the average population

exposure and by accounting for the variability within the core city that gives rise to higher exposure for a subset of the population, here quantified as the 95%-percentile. For $PM_{2.5}$, the current proxy used for residential heating is not sufficiently accurate that downscaling $PM_{2.5}$ gives significant improvement, and/or the gridding of primary PM in the original 0.1x0.1 degree emission data set is not of high enough quality.

The relative increase from downscaling in traffic contribution was shown to be very variable in time, due to changing meteorological conditions, so a constant factor does not seem appropriate. One option is to run the uEMEP model operationally for the 79 cities and use its outputs to calculate the population-weighted mean and 95-% percentile of exposure for each hour, as done in this study. The computational cost of uEMEP increases strongly with resolution, but at 250 m resolution it is manageable to run downscaling of all the cities for the 4-day forecast horizon within the CAMS2_71 Service as it is set up at present. This would need to be done after the Baseline EMEP calculation has completed, whose output must include the local fractions. Based on the computation times of the runs of this study, all cities can then be run within 1 hour on a computational node with 96 parallel CPUs, each with 2 GB memory. If a coarser resolution of 500 m or 1 km was used instead, the computational cost would be lower.

A possible visualization of the resolution effect could be based on Figure 4.4a, but removing the 5-95% percentile range for EMEP.

Ideally, the modified population-weighted mean concentration due to downscaling should also appear in the “City & Country impacts” page. It might seem natural that the increase in mean concentration due to downscaling should be attributed in full to the city's contribution to itself. However, it is actually not so straightforward, since the downscaling domain is not defined by the city boundary, but instead a moving window of size 2x2 EMEP cells around every receptor grid cell. 2x2 EMEP cells correspond approximately to a 30x30 km area. Depending on the location of the receptor, a smaller or larger part of this area is outside the core city. Sources outside the core city should not be attributed to the city. However, uEMEP has newly implemented a feature where the contribution from inside vs. outside a polygon can be separately quantified. This feature was not used in the runs done for this study, but it could be used to separate the increase from downscaling into an increase from sources inside the core city polygon vs. outside. The increase due to sources outside the core city could be attributed to the “Rest of country” source. This still has two important limitations, since the downscaling categorizes sources by GNFR sector and not by country: 1) If a neighboring country is close enough to the city that emission sources from that country falls inside the downscaling domain, any increase from such sources due to downscaling will be attributed to “Rest of country” instead. 2) International shipping is shown as a separate source in the service; however, uEMEP does not distinguish domestic and international shipping, so changes in contributions from international shipping due to downscaling will instead be attributed to the city’s contribution to itself or “Rest of country”, depending on the emission location.

4.9 References

Denby, B. R., Gauss, M., Wind, P., Mu, Q., Wærsted, E. G., Fagerli, H., Valdebenito, A., and Klein, H.: Description of the uEMEP_v5 downscaling approach for the EMEP MSC-W chemistry transport model, Geosci. Model Dev., 13, 6303–6323, https://doi.org/10.5194/gmd-13-6303-2020, 2020.

EMEP status report 1/2023, https://emep.int/publ/reports/2023/EMEP_Status_Report_1_2023.pdf

Jalkanen, J.-P., Brink, A., Kalli, J., Pettersson, H., Kukkonen, J., Stipa, T., 2009. A modelling system for the exhaust emissions of marine traffic and its application in the Baltic Sea area. Atmos. Chem. Phys. 9, http://dx.doi.org/10.5194/acpd-9-15339-2009.

Kystverket: AIS global shipping emission data, directly provided by Kystverket to MET Norway, available at: https://www.kystverket.no/en (last access: 11 January 2022), 2020.

Hassani, Amirhossein; Salamalikis, Vasileios; Schneider, Philipp; Stebel, Kerstin; Castell, Núria (2025). Harmonized, Standardized, and Corrected Crowd-Sourced Low-Cost Sensor PM2.5 Data from Sensor.community and PurpleAir Networks Across Europe. figshare. Dataset. https://doi.org/10.6084/m9.figshare.27195720.v1

Mu, Q., Denby, B.R., Wærsted, E.G., Fagerli, H., 2022. Downscaling of air pollutants in Europe using uEMEP_v6. Geosci. Model Dev. 15 (2), 449–465. http://dx.doi.org/10. 5194/gmd-15-449-2022, URL: https://gmd.copernicus.org/articles/15/449/2022/.

OpenStreetMap contributors, 2020. Planet dump. retrieved from https://planet.osm.org, URL: https://www.openstreetmap.org.

Pesaresi, M., Schiavina, M., Politis, P., Freire, S., Krasnodębska, K., Uhl, J. H., et al. (2024). Advances on the Global Human Settlement Layer by joint assessment of Earth Observation and population survey data. International Journal of Digital Earth, 17(1). https://doi.org/10.1080/17538947.2024.2390454 (population density data retrieved from: https://human-settlement.emergency.copernicus.eu/download.php?ds=pop)

Schiavina, M., Melchiorri, M., Pesaresi, M., Politis, P., Carneiro Freire, S.M., Maffenini, L., Florio, P., Ehrlich, D., Goch, K., Carioli, A., Uhl, J., Tommasi, P. And Kemper, T. (2023), GHSL Data Package 2023, Publications Office of the European Union, Luxembourg, 2023, doi:10.2760/098587, JRC133256 .

Wind, P., Rolstad Denby, B., and Gauss, M. (2020): Local fractions – a method for the calculation of local source contributions to air pollution, illustrated by examples using the EMEP MSC-W model (rv4_33), Geosci. Model Dev., 13, 1623–1634, https://doi.org/10.5194/gmd-13-1623-2020.

5 Propagation of emission uncertainties into the city SR service and potential for operational implementation

5.1 Introduction

Providing uncertainty quantification (UQ) estimates alongside air quality forecasts is crucial to maximizing their utility. Uncertainties arise from, among others, uncertainties in emissions, meteorology, model uncertainties, resolution of the model and emissions, and city definitions. Here we focus on the propagation of uncertainties in emissions and temporal profiles.

Commonly employed methods for propagating uncertainty through chemical transport models (CTMs) include single parameter perturbations, ensemble methods and Monte Carlo techniques (e.g., Sax and Isakov, 2003, Aleksankina et al., 2018), and emulator (also known as “surrogate”) modeling approaches (e.g., Aleksankina et al., 2019, Guo et al., 2024).

Monte Carlo techniques are common tools for analyzing uncertainties arising from complex models. In such an approach with a CTM, one could sample from the prescribed emissions distributions, use these samples as input to run the model, and estimate the uncertainty of the prediction by considering the prediction spread across the ensemble. Hence, each sample in the input space yields a corresponding model run. The compute requirements of modern CTMs yield this method computationally prohibitive in many applications, and in particular unsuitable for operational air quality forecasts, such as those provided by CAMS.

Emulator models serve as another common alternative to either estimate or propagate uncertainties through complex models. They can be employed either alongside Monte Carlo techniques or as a substitute for them (e.g., Conibear et al., 2021). Used with Monte Carlo techniques, the aim is that cheaper evaluation of the emulators allows efficient exploration of the parameter space and that this yields a more complete understanding of the prediction uncertainty. Using emulator models alone relies on using an emulator model which provides a more complete understanding of the predictive distribution, and assuming this predictive distribution is representative of the uncertainty in the original CTM.

The most mathematically complete description of prediction uncertainty available would be the entire predictive distribution. In some cases, this is obtainable if one has a formula describing a system and they wish to propagate uncertainty through this formula. CTMs are highly complex, and therefore no single formula is available to describe the entire system. This does not mean, however, that analytical methods to propagate emissions uncertainties are not available. For example, previous work by Dunker et al. (2022) has combined analytical methods with Monte Carlo techniques to estimate and analyze uncertainties in ozone predictions arising from biogenic and anthropogenic emissions. Given a method to differentiate the model, analytical approximations for uncertainty propagation may be used.

Utilizing analytical methods is preferable for a few reasons. For one, compared to emulator model-based techniques, we avoid an additional modeling step because we do not need to train another model. Furthermore, we do not need to make an additional assumption that our emulator model correctly characterizes both the first-order and higher-order properties of the predictive distribution. Analytical methods can be implemented efficiently because they provide a formulaic representation of the uncertainty. Moreover, they are relatively explainable and easy to understand, with complete information available regarding where sources of uncertainty arise from, and how they interact. Finally, using analytical methods for uncertainty propagation allow modelers the ability to make modeling choices, balancing accounting for important sources of uncertainty and computational efficiency. Modelers have complete control in this aspect, and therefore can gain a more complete understanding of how emissions uncertainties result in forecast uncertainties.

The objective of this chapter is to explore an approach to propagate (yearly country, sector, emitted component and temporal) emission uncertainties through the EMEP MSC-W

(European Monitoring and Evaluation Programme Meteorological Synthesizing Centre-West) model with direct applications to operational air quality forecasts provided by CAMS.

5.2 Emissions Uncertainties

Emissions uncertainties are provided by Work Package 5 in CAMEO, and described in the deliverables Super and Kuenen [2023] and Super et al. [2024]. Uncertainties are provided as yearly estimates in terms of absolute and relative standard deviations, per country, sector, emitted component, and in the case of primary particulate matter (PPM), as a gridded data product matching the model resolution of 0.1° x 0.1°. Following the recommendations of the data providers we work exclusively with the relative standard deviations[4]. A general first analysis can be made by cross-referencing the contribution and uncertainty data.

Table 5.1: The mean relative standard deviations of primary $PM_{2.5}$ emissions for GNFR sectors across countries.

Sector	**Mean Relative Standard Deviation**
PublicPower	0.75
Industry	0.39
OtherStationaryComb	0.33
Fugitives	0.47
Solvents	0.56
RoadTransport	0.08
Shipping	0.35
Aviation	0.31
OffRoad	0.24
Waste	0.93
AgriLivestock	0.21
AgriOther	0.21

Emissions uncertainties at the sector level are inherently relevant insofar as each sector contributes to anthropogenic air pollution. The following figure provides an illustration for how

[4] The relative standard deviation, also known as the coefficient of variation, is a way to express the standard deviation as a percentage of the mean. It provides a sense of the variability in the emissions relative to the average.

each sector proportionally contributes to the anthropogenic part of $PM_{2.5}$ concentrations across seasons in 2019 (calculated with the Local Fraction (LF) methodology). However, the relative sector contribution to $PM_{2.5}$ concentrations varies widely from country to country[5]. We additionally may aggregate across the entire year, considering maps of the relative contribution from each emission sector to $PM_{2.5}$ concentrations to get a sense of how different sectors proportionally contribute at different locations, shown in Figure 5.2. Seasonal versions of these plots are available in Appendix B. Generally we find that the anthropogenic part of $PM_{2.5}$ in summer is less certain in a relative sense than winter. Part of the reason for this is that industry and public power have higher average relative standard deviations and contribute more in the summer. Another potentially contributing factor is that road transport contributes relatively much more in the winter than in summer. Compared to other sectors, road transport is very certain with a sector standard deviation of 0.08 (the lowest of all sectors). This has the effect of exacerbating the difference between winter and summer to make winter appear relatively more certain. Agricultural livestock also has a relatively low relative standard deviation (0.21) and contributes much more in the winter, potentially exacerbating the seasonal differences.

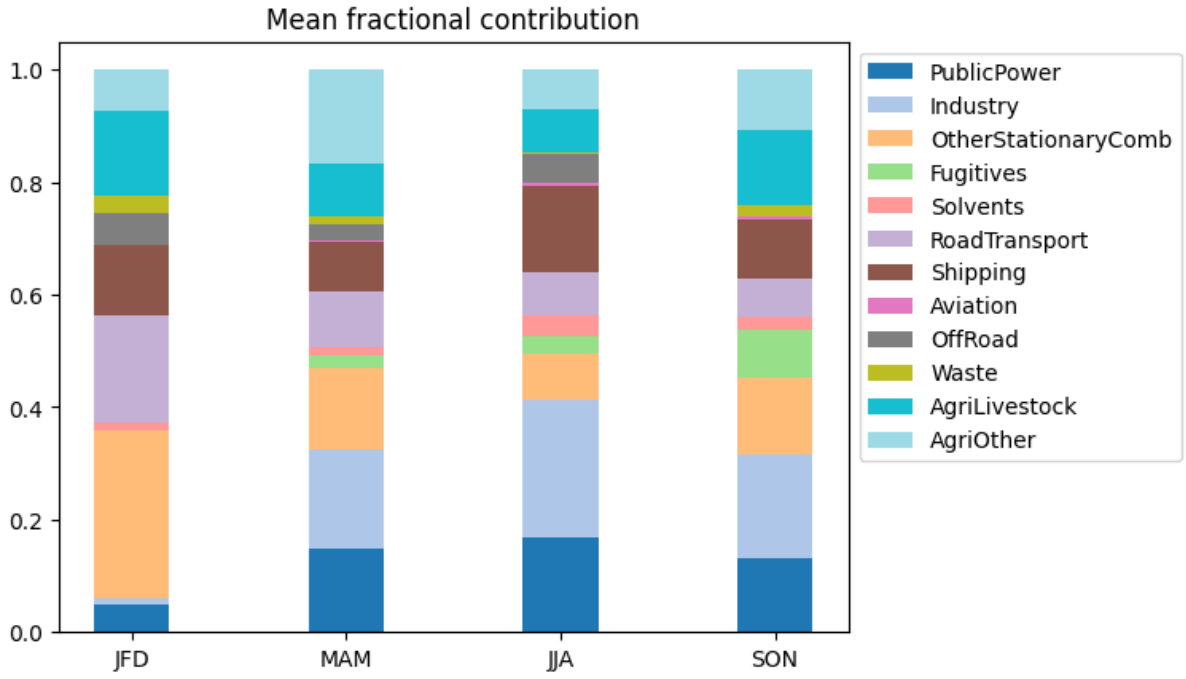


Figure 5.1: Mean fractional contribution to the anthropogenic part of the $PM_{2.5}$ concentrations across GNFR sectors during seasons in 2019 over the EMEP domain (both sea and land).

[5] A table of countries from highest to lowest relative standard deviations is provided in the appendix.

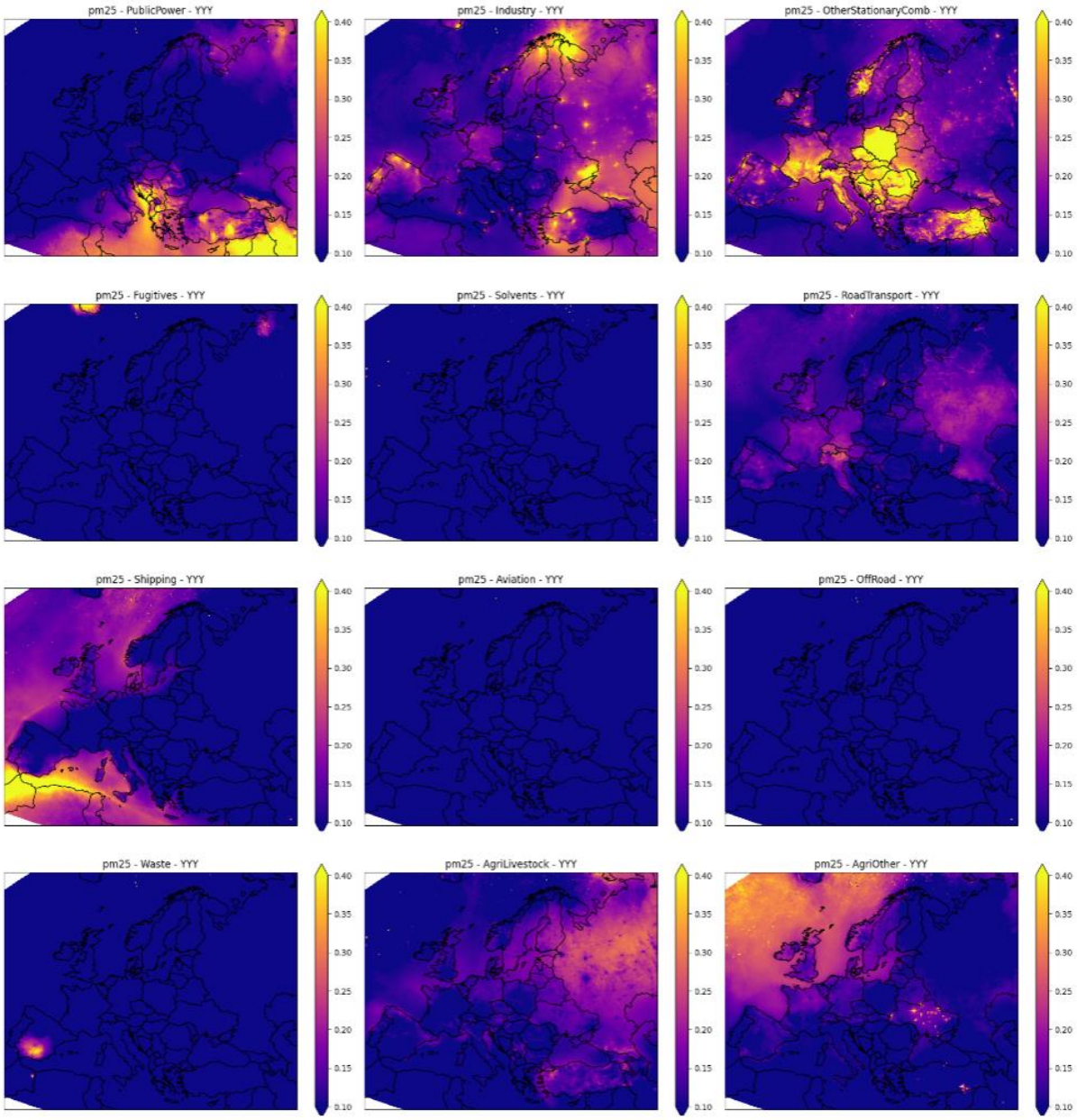


Figure 5.2: The contribution fraction from each sector to anthropogenic $PM_{2.5}$ concentrations in 2019. Scale is from 10% - 40%.

5.3 An Analytical Method for Propagation of Emission and Temporal Uncertainty

We combine the emission uncertainties provided by TNO in conjunction with sensibilities to emission changes obtained using the Local Fraction method (LF) (Wind et al., 2020, 2025). Using these as inputs, we propagate the uncertainty forward using a probabilistic approach based on a local expansion of the EMEP model as a first-order Taylor series. The key insight to this method is that the local fractions output can be seen as the constituent terms in the Taylor expansion of the pollutant concentrations expressed as functions of emissions. In

conjunction with the emissions uncertainties, we derived uncertainties in the EMEP output using the variance formula (Cramér, 1999, Ku et al., 1966).

Let $\mathcal{L}$ represent the set of countries, $\mathcal{S}$ the set of GNFR sectors, and $\mathcal{E}$ the set of emitted components. Specifically, the set of emitted components considered is:

$$\mathcal{E} = \{PPM, SO_x, VOC, NO_x, NH_3\}$$

For each combination of country, sector, and emitted component (i.e., for all $\boldsymbol{c} \in \mathcal{L} \times \mathcal{S} \times \mathcal{E}$, denoted as $\mathcal{C}$), we have estimates for both the emission mean (E_c) and the relative standard deviation (σ_c^{rel}), where: $\sigma_c^{rel} = \sigma_c / E_c$.[6]

Let C represent the pollutant concentration. We write the LF output in the following way for ease of notation:

$$LF_{\mathbf{c}} := \frac{\partial C}{\partial E_{\mathbf{c}}} E_{\mathbf{c}}.$$

The Local Fractions relate the amplitude of small variations of emissions to the resulting variations of pollutant concentrations. We define the total air pollution accounted for as the sum of the contributions over all $\boldsymbol{c} \in \mathcal{C}$:

$$P_{tot} = \sum_{\mathbf{c}\in\mathcal{C}} LF_{\mathbf{c}} = \sum_{\mathbf{c}\in\mathcal{C}} \frac{\partial C}{\partial E_{\mathbf{c}}} E_{\mathbf{c}}$$

We can express the uncertainty in P_{tot} by accounting for the uncertainties attributed to each of the contributions.

$$\mathrm{Var}(P_{tot}) = \mathrm{Var}\left(\sum_{\mathbf{c}\in\mathcal{C}} LF_{\mathbf{c}}\right)$$

Using the properties of variance, this can be expanded in the following manner.

$$\mathrm{Var}\left(\sum_{\mathbf{c}\in\mathcal{C}} LF_{\mathbf{c}}\right) = \sum_{\mathbf{c}\in\mathcal{C}} \mathrm{Var}(LF_{\mathbf{c}}) + \sum_{\mathbf{c}\in\mathcal{C}} \sum_{\substack{\mathbf{c}'\in\mathcal{C} \\ \mathbf{c}'\neq\mathbf{c}}} \mathrm{Cov}(LF_{\mathbf{c}}, LF_{\mathbf{c}'}).$$

Let us consider the terms on the right-hand side of this equation. The first term can be written as

$$\sum_{\mathbf{c}\in\mathcal{C}} \mathrm{Var}(LF_{\mathbf{c}}) = \sum_{\mathbf{c}\in\mathcal{C}} \sigma_{\mathbf{c}}^2$$

where

$$\sigma_{\mathbf{c}} \approx LF_{\mathbf{c}} \cdot \sigma_{\mathbf{c}}^{Rel}$$

Mathematical formalism for what the LF output represents can be found in (Wind and van Caspel, 2025, Equations 3, 4). This final approximation is known entirely as it is the product of the LF output and the relative standard deviation for $\boldsymbol{c} \in \mathcal{C}$.

[6] Note that while we do have absolute standard deviations provided by TNO, they refer to a specific emissions data set, and therefore are not directly transferable to other emissions data sets. We therefore assume that, while the absolute emissions and standard deviations can change between the provided data and the emissions data used in our EMEP model runs, the relative standard deviations do not change, and so we directly work with just the relative standard deviations.

Inspecting the second term in the sum on the right-hand side of the variance formula, we can momentarily consider the covariance. We can decompose the formula for the covariance

$$\mathrm{Cov}(LF_{\mathbf{c}}, LF_{\mathbf{c}'}) = \sigma_{\mathbf{c}}\sigma_{\mathbf{c}'}\rho(LF_{\mathbf{c}}, LF_{\mathbf{c}'}),$$

where $\rho(\cdot,\cdot)$ is the correlation. The correlations are approximated using the formulation above. For the correlation term, we have some options for simplifying assumptions.

1. Uncorrelated case: Assume $\rho(LF_{\boldsymbol{c}}, LF_{\boldsymbol{c}'}) = 0$ for all $\boldsymbol{c}, \boldsymbol{c}' \in \mathcal{C}$.
2. Fully correlated case: Assume $\rho(LF_{\boldsymbol{c}}, LF_{\boldsymbol{c}'}) = 1$ for all $\boldsymbol{c}, \boldsymbol{c}' \in \mathcal{C}$.
3. Sector correlated case:

$$\rho(LF_{\mathbf{c}}, LF'_{\mathbf{c}}) = \begin{cases} 1 & \mathbf{c}, \mathbf{c}' \text{share the same sector} \\ 0 & \text{otherwise} \end{cases}$$

Each of these correlation cases relates to specific assumptions about emission correlations. In the first case, assuming zero correlation implies that no emission uncertainty is informative in predicting any other emissions uncertainties. For example, the uncertainty in emissions produced by residential heating in Norway would tell us nothing about the uncertainty in emissions from the same source in Sweden, despite the fact that these countries are closely geographically related and likely experience similar meteorological conditions driving these emissions. In the second case, we consider the exact opposite scenario, which is also very unrealistic. In such a situation, the random variable describing emissions uncertainty from Fugitives in Montenegro would be 100% correlated, with the random variable describing emissions uncertainty from residential heating in Sweden. Likewise for all other combinations of uncertainty in emissions in all countries, sectors, and emitted components. This gives a maximalist approximation on the variance. Finally, in case three, we consider emissions fully correlated within sectors, but not across sectors. For example, uncertainty in NO_x emissions from the transportation sector in the Netherlands would be perfectly correlated with uncertainty in NO_x emissions in Hungary, but uncertainty in emissions from NO_x in the Dutch transportation sector would be uncorrelated emissions from all other sectors in the Netherlands. In this sense, while we perhaps overestimate the correlation within sectors, we aim that this approximation can compensate for some of the missing correlations within countries.

There are some computational considerations pertaining to these cases, which make them attractive approximations to consider. In case 1, the variances simply add. This places the most conservative estimate on the prediction variance we consider and can be interpreted as a lower bound. In case 2, this assumption actually allows us to simply add the standard deviations[7]. Case three, where we consider all sectors fully correlated with themselves and assume emissions uncertainties between sectors are uncorrelated, represents one compromise between these two extremes. In the sector-correlated case, it suffices to add the squares of the sums of the standard deviation within each sector, which also yields computational benefits.

One important consideration is how assumptions about correlations affect our results. In the uncorrelated case, a well-known statistical principle[8] justifies the use of a normal (Gaussian) error distribution. However, this justification may break down when inputs to the model exhibit correlation. When sectors are independent of each other, it's standard to assume that model errors also follow a normal distribution. However, if sectors are correlated, this assumption might not hold as well. We do not assume strong correlations between sectors, so it's still reasonable to use the normal distribution as an approximation even when combining data from thirteen sectors. We keep this assumption throughout the analysis because it keeps the methods consistent and enables closed-form analytical solutions. This approach also allows

[7] This is justified because $a^2 + 2ab + b^2 = (a + b)^2$, which implies that the standard deviation can be written simply in terms of a + b.

[8] The Central Limit Theorem justifies the use of a Gaussian error distribution.

us to calculate prediction intervals, which help us compare how well the model performs under different correlation assumptions.

More realistic correlation structures were also considered. Based on the experience creating the uncertainty data product in CAMEO Work Package 5, TNO provided a first approximation of how the uncertainties might be correlated. Table 5.2 presents these suggestions. Implementation of such a correlation structure is significantly more complicated, however, from both purely a technical and computational perspective. One of the key features of the aforementioned correlation structures is their computational benefit, which makes them not only computationally feasible but potentially also suitable for production services. As we consider larger and more complex correlation structures, their implementation can quickly become computationally prohibitive. For this reason, we do not implement this suggested correlation structure, however this could be attempted in the future.

Table 5.2 Proposed correlation structure by TNO for emissions uncertainties (how much errors in e.g. GNFR A in one country is correlated with errors in GNFR A in another country). Percentages represent the share of each sector to total emissions.

GNFR	GNFR Name	$PM_{2.5}$	NOX	NMVOC	SOX	NH3	Expected corr. in error ranges	Explanation
A	Power plants	2%	14%	1%	29%	0%	0	Point sources, reporting by facility in many countries
B	Industry	13%	20%	9%	42%	1%	0.25	Point sources, reporting by facility in many countries - but also more diffuse industrial sources, which could be using similar (Guidebook) methods across countries
C	Residential	64%	12%	15%	19%	1%	0.5	Mix between countries using their own methodology vs. countries using the EMEP/EEA Guidebook; take an average
D	Fugitives	1%	0%	6%	9%	0%	0.25	Small sector, different subsectors and methods for estimating emissions, likely low correlations
E	Solvents/product use	2%	0%	38%	0%	2%	0.75	Very uncertain category, many countries use default methods. Also countries may incorporate (parts of) methods from other countries for themselves
F	Road transport	9%	42%	7%	0%	1%	0.75	Many countries use COPERT (~75% of EU countries), some use other methods (e.g., HBEFA), but these are also compared sometimes. So likely correlations do exist to a large extent, but not fully.
G	Shipping (inland)	0%	1%	0%	0%	0%	0	This concerns only inland shipping where correlations are probably quite low since countries have their own methods.
H	Aviation	0%	1%	0%	0%	0%	1	Expect full correlations since most approaches rely on the ICAO data and Eurocontrol approach. So all using the same type of data in general.
I	Other mobile	2%	9%	2%	0%	0%	0.25	Methods differ quite a lot between countries, also many subsectors included here with all their own characteristics.
J	Waste	4%	0%	1%	0%	2%	0.5	Also different subsectors. Not very important for AQ, so probably not that relevant anyhow.
K	Agriculture livestock	2%	0%	18%	0%	42%	0.75	Assume largely using the Guidebook (most countries) with exception of some countries with important agriculture sector (e.g. NL, DE, FR)
L	Agriculture other	1%	0%	3%	0%	51%	0.25	Methods differ quite a lot between countries, also many subsectors included here with all their own characteristics.
G	Shipping (sea)						1	Generally one method is used across all sea regions, so full correlation

5.4 Total Uncertainty Arising from Anthropogenic Emissions Uncertainty assuming Sector Correlation

It will be useful to write explicitly the total uncertainty arising from anthropogenic emissions assuming the sector correlated case. This correlation structure offers some realism while also facilitating computation benefits, which we explicitly write below.

Fix time t. Let $P_{tot}(\cdot\ ,\ t)$ represent the spatial field of total accounted for air pollution at time t. Then we have:

$$P_{tot}(\cdot, t) = \sum_{s \in \mathcal{S}} LF(\cdot, t, s)$$

Assuming that s is uncorrelated with s', for all $s, s' \in \mathcal{S}$,

$$\mathrm{Var}\left(P_{tot}(\cdot, t)\right) = \sum_{s \in \mathcal{S}} \mathrm{Var}\left(LF(\cdot, t, s)\right)$$

$$LF(\cdot, t, s) = \sum_{\ell \in \mathcal{L}} LF(\cdot, t, s, \ell)$$

Assuming full correlation within sector s across countries:

$$\mathrm{Var}\left(LF(\cdot, t, s)\right) = \mathrm{Var}\left(\sum_{\ell \in \mathcal{L}} LF(\cdot, t, s, \ell)\right) = \left(\sum_{\ell \in \mathcal{L}} \sigma(\cdot, t, s, \ell)\right)^2.$$

Assuming full correlation within sector s across emitted components:

$$\sigma(\cdot, t, s, \ell) = \sum_{e \in \mathcal{E}} \sigma(\cdot, t, s, \ell, e)$$

$$\sigma(\cdot, t, s, \ell, e) \approx \sigma^{Rel}(s, \ell, e) LF(\cdot, t, s, \ell, e)$$

5.5 Maps Exploring Correlation Structures

We now explore these various correlation structures as both yearly averages and across seasons to determine how their associated assumptions yield various relative uncertainty maps. Doing so allows modelers to gauge approximate ranges and their plausibility. We now present maps of the relative standard deviations induced by the three considered correlation structures.

Generally speaking, the uncorrelated case yields uncertainty maps which are considered too small. This has the implication that a prediction interval data product derived from a predictive distribution assuming no correlation between the inputs would underestimate the actual uncertainty and not adequately capture the spread in observations. Considering the other extreme, assuming full correlation between all inputs yields uncertainty maps which are generally too high. This would result in prediction intervals which are too large, and not provide meaningful uncertainties alongside air quality forecasts. Assuming purely sector correlation balances these extremes, in a manner which yields more realistic uncertainty maps, which are more consistent with the spread in observations and domain expertise.

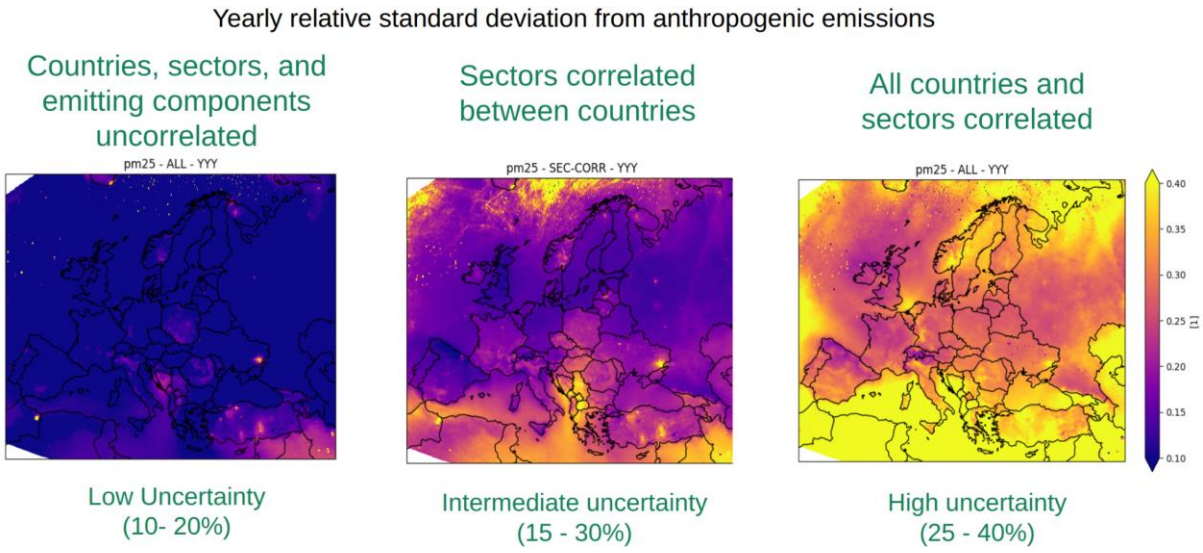


Figure 5.3: Maps of the yearly relative standard deviation of $PM_{2.5}$ *in 2019 from anthropogenic emissions across different assumed correlation structures. The scale ranges from 10% - 40%.*

5.6 Formulas for Estimating Local Contribution Uncertainty Using Gridded Uncertainty Data

In CAMS, it is often useful to consider the local contribution to air pollution concentrations, for example of $PM_{2.5}$ (defined here as the contribution from a city to itself). It is therefore important to also consider the uncertainty in the local contribution to the concentrations. To motivate with an example, we will consider a specific air pollution episode beginning 16th of February 2019 in Bratislava. We will explore this example in more detail later in this chapter, however for now we simply use it as a concrete example to illustrate how we can account for uncertainty arising from local emissions.

Let $C_{PM2.5}$ represent the surface concentration of $PM_{2.5}$. Furthermore, let D represent the domain. In this example we have that $|D| = 400 \times 260$. For fixed $\mathbf{x} \in D$ and time t, let $C_{PM2.5}(\mathbf{x}, t)$ be the surface concentration of $PM_{2.5}$ at the given spatio-temporal coordinates. We may hold out one of these coordinates as generic, in which case a $\cdot$ will represent that this is to be taken over all values of the held out coordinate. For example, $C_{PM2.5}(\cdot, t) \in \mathbb{R}^{|D|}$ is the spatial field at time t, whereas $C_{PM2.5}(\mathbf{x}, \cdot)$ represents the time series of concentrations of $PM_{2.5}$ at location $\mathbf{x}$.

Let B denote a set of cities. Fix $b = Bratislava \in B$. Let $\mathbf{m}_b \in \mathbb{R}^{|D|}$ denote the city core mask for Bratislava (see definition of city in Chapter 3.1). Each entry of this array is the fraction of each grid cell which is defined as belonging to Bratislava by the core-city mask definition. For a given point in time t, the background concentration in Bratislava is given by

$$C_{\mathrm{PM2.5}}(b,t) = \mathrm{vec}(\mathbf{m}_b)^\top \mathrm{vec}\left(C_{\mathrm{PM2.5}}(\cdot,t)\right).$$

We would like to express the uncertainty (i.e., the variance) of this quantity, ${\sigma_b}^2(t)$. However, we do not have uncertainties for all contributing sources. Rather, we write the concentration in the following way

$$C_{\mathrm{PM2.5}}(\cdot,t) = P_{tot}(\cdot,t) + \epsilon(\cdot,t),$$

where the closure term $\epsilon(\cdot, t)$ is what remains between all anthropogenic sources accounted for as defined above and the total concentration. It therefore accounts for natural sources and nonlinear chemistry. We lack uncertainties for this, and therefore we only consider the uncertainty in P_{tot}.

$$P_{tot}(b,t) = \sum_{s \in \mathcal{S}} LF(b,t,s)$$

where $LF(b,t,s)$ is the contribution from sector s to the concentration of $PM_{2.5}$. Specifically

$$LF(b,t,s) = \text{vec}\left(\mathbf{m}_b\right)^\top \text{vec}\left(LF(\cdot,t,s)\right)$$

.

Alternatively we could write

$$P_{tot}(b,t,s) = \sum_{\mathbf{x} \in b} \mathbf{m}_b(\mathbf{x}) LF(\mathbf{x},s,\text{Slovakia})$$

$$LF(\mathbf{x},s,\text{Slovakia}) = \sum_{e \in \mathcal{E}} LF(\mathbf{x},s,\text{Slovakia},e)$$

.

Assuming full correlation within sector s across emitted components, we can compute their variance as the sum

$$\text{Var}\left(\sum_{e \in \mathcal{E}} LF(\mathbf{x},s,\text{Slovakia},e)\right) = \left(\sum_{e \in \mathcal{E}} \sigma(\mathbf{x},t,s,\text{Slovakia},e)\right)^2,$$

where each of $\sigma(\mathbf{x},t,s,Slovakia,e)$ is specific to the emitted component. We make this choice because we have a gridded data product specifically available for primary PM (PPM), but not other emitted components, and it is desirable to take full advantage of this gridded data product, particularly when considering the uncertainty arising from local emissions.

The Local Fractions methodology also provides us a way to determine for each grid cell $\mathbf{x}$, the neighboring cells (within a window), the amount of air pollution which comes from that neighboring cell. The collection of all neighboring cells, we will call a "Local Fractions Window", and denote as $LFW(\mathbf{x})$. It is important to note, however, that not all air pollution which reaches location $\mathbf{x}$ comes from $LFW(\mathbf{x})$, rather only a fraction. The fraction of air pollution at location $\mathbf{x}$ which comes from the k^{th} neighbouring cell is denoted as $lf_k(\mathbf{x})$.

For all emitted components except for primary PM, we calculate $\sigma(\mathbf{x},t,s,Slovakia,e)$ the same way as before, by multiplying the Local Fractions output for that country, sector, emitted component and multiplying by its associated relative standard deviation. In the case of PPM, however, we wish to use the gridded uncertainty data product for all neighboring cells in $LFW(\mathbf{x})$ for the proportion of air pollution which reaches $\mathbf{x}$ from $LFW(\mathbf{x})$, and balance that with the standard definition for all other emitted components for the remaining proportion. For simplicity, we assume these two sources are uncorrelated, and add their variances. The complete definition for $\sigma(\mathbf{x},t,s,Slovakia,e)$ can therefore be given as

$$\sigma(\mathbf{x},t,s,\text{Slovakia},e) = \begin{cases} \left[\sigma^2_{LFW}(\mathbf{x}) + \left(1 - \sum_k lf_k(\mathbf{x})\right)\left(\sigma^{Rel}_{\text{Slovakia},s,\text{PPM}} LF(\mathbf{x})\right)^2\right]^{\frac{1}{2}}, & \text{if } e = \text{PPM} \\ \sigma^{Rel}_{\text{Slovakia},s,e} LF(\mathbf{x}) & \text{otherwise} \end{cases}$$

$$\sigma^2_{LFW}(\mathbf{x}) = \sum_{i=1}^{|LFW(\mathbf{x})|} \left(LFW(\mathbf{x})_i \sigma_i^{Rel}\right)^2 + 2 \sum_{\substack{i,j \\ j>i}} \text{Cov}(LFW(\mathbf{x})_i, LFW(\mathbf{x})_j),$$

$$\mathrm{Cov}(LFW(\mathbf{x})_i, LFW(\mathbf{x})_j) = \sigma_i \sigma_j \exp\left(\frac{-d(i,j)}{L_{\mathrm{Slovakia}}}\right),$$

$$\sigma_i = \sigma^{Rel}_{s,\mathrm{Slovakia,PPM}}(i) LFW(\mathbf{x})_i.$$

Note that to use the gridded PPM data product in each Local Fraction Window, we need to account for the spatial correlation between the cells. This correlation is accounted for by the spatial covariance function, which is specified to be an exponential. The range parameter (here $L_{Slovakia}$) is provided for each country, so that the aggregated gridded uncertainties for PPM match the country sector level uncertainties provided in Super and Kuenen [2023]. The procedure by which this is accomplished is described in Super et al. [2024]. The distance metric used in the great-circle distance, although on such small scales such as in our application, this is essentially the same as the Euclidean distance.

5.7 Uncertainties in Emission Temporal Profiles

Emissions are reported by countries per sector at an annual frequency. As a part of their work for CAMS, TNO and BSC distribute these emissions throughout time and space for each country, such that the accumulated emissions are equal to the annual reported values. Typically this involves the use of “time factors” in the CTMs, which represent the fraction of the emissions that should be assigned to each category. Commonly employed time factors are monthly, day-of-week, and hour-of-day time factors. Interpolation and reweighting and procedures are then applied to remove large discontinuities in the time factor series, and to ensure consistency.

We would additionally like to account for the uncertainty which arises from how the emissions are distributed throughout time. Uncertainties in temporal profiles have been produced and are described in Guevara et al. (2024). According to the documentation, these temporal profiles are pertinent for characterizing the uncertainty associated with how modelers may implement them within their modeling frameworks.

We account for this source of uncertainty in the following manner. Consider the time factor for a given point in time at an hourly resolution, for a given pollutant, country, and sector. Ignoring an additional interpolation step, the time factors used are as follows:

$$\sum f_M = 12, \sum f_W = 7, \sum f_H = 24.$$

The data accompanying deliverable 5.1 described in Guevara et al. [2024] provides estimates of the means and standard deviations for each of these summands. Let $f_M f_W f_H =: f_t$. Assuming independence of the time factors and applying properties of expectation and variance, we have that $\mathbb{E}f_t = \mathbb{E}f_M \mathbb{E}f_W \mathbb{E}f_H$, and

$$\mathbb{V}ar(f_t) = (\mathbb{V}ar(f_M) + (\mathbb{E}f_M)^2)(\mathbb{V}ar(f_W) + (\mathbb{E}f_W)^2)(\mathbb{V}ar(f_H) + (\mathbb{E}f_H)^2) - (\mathbb{E}f_M \mathbb{E}f_W \mathbb{E}f_H)^2.$$

Hence, for each country, sector, and emitted component, at every time step t, we have an estimate for the variance of the associated time factor. Note that while the formula we derive is exact based on the independence assumption of the time factors, the data we use are variance estimates, and so the resulting variance is itself an estimate.

We then use this variance and the mean estimate to derive a relative standard deviation for each country, sector, and emitted component. By multiplying this with the LF output, we have an estimate of the standard deviation which arises from the temporal profile. The LF methodology does not have native way to introduce this source of temporal uncertainty, so rather we consider it an additional source of uncertainty which is uncorrelated from emissions uncertainty, and simply add it on to our uncertainty estimate. However, we can not assume

that the temporal uncertainty at time t manifest instantaneously into the concentration uncertainty at time t. Rather, some form of aggregation method is needed, the various methods of which are explored in subsequent sections.

We summarize the inclusion of the temporal uncertainty with the following procedure:

1. Use the country, sector, emitted component temporal uncertainties from Guevara et al. [2024].
2. For every time step, compute the time fraction (mean) and its standard deviation assuming independence.
3. Compute a relative standard deviation, and then multiply this relative standard deviation by the associated Local Fractions output for that country, sector, emitted component.
    a. We can think of this as the instantaneous uncertainty in units of the data arising from the temporal profile.
4. Aggregate all of these uncertainties, assuming that the temporal profiles are 100% correlated across sectors
    a. This correlation structure need not be the same as the emissions
5. Consider an aggregated uncertainty for that day (e.g., the max of the previous 24 hour period, or a 24 hour rolling window).
6. Add on this temporal variance to the emission variance.

5.8 Analysis of air pollution episodes

It is desirable to assess the effectiveness of the method for propagating uncertainties in emission amount and temporal profiles under a selection of pollution episodes caused by different emission sources. To do so, we choose the following air pollution episodes during 2019 in European cities to gauge the method's performance. For example, when an episode is primarily driven by emissions in a country or sector which is fairly uncertain, we anticipate that the resulting uncertainty will be larger and perhaps capture the spread of most of the observations (and conversely if the episode is primarily driven by emissions in a fairly certain country or sector). The chosen episodes may also be analyzed using the currently available tools on the CAMS policy service website: https://policy.atmosphere.copernicus.eu/yearly/

Table 5.3: Air pollution episodes used to investigate the propagation of emissions uncertainty method.

City	Time	Type of episode
Bratislava	2019.02.12 - 2019.02.18	$PM_{2.5}$ arising from emissions in residential heating and industry
Frankfurt	2019.02.12 - 2019.02.19	$PM_{2.5}$ arising from emissions in residential heating and industry
Amsterdam	2019.03.16 - 2019.03.25	$PM_{2.5}$ arising from emissions in the transportation and agricultural sectors.
Warsaw	2019.10.19 - 2019.10.24	$PM_{2.5}$ arising from industry, traffic and agriculture

For each of the following episodes, we provide figures of the concentrations broken down by the Local Fractions output into the sector, country, and emitted component. Doing so allows

analysis of how each of the various sources contributes while integrating out the other dimensions (e.g., analyzing how the sectors contribute to the $PM_{2.5}$ concentrations, while aggregating over the emitted components in all of the contributing countries).

We additionally provide a prediction interval data product derived using our method compared with observations. In these figures, we show the pollutant concentrations with the prediction intervals derived from our method to propagate the emissions and temporal uncertainties and an assumed Gaussian distribution of the errors[9]. The 95% prediction intervals are also decomposed into the portion arising from local emissions, all emissions, and the temporal uncertainty. In all of the following figures, a rolling 24 mean of the temporal uncertainty is used, however this is subject to further exploration in Section 5.8.5. Observations are also incorporated into the analysis, categorized by station type (i.e., suburban or urban) to enable the calculation of empirical coverage. While data from both urban and suburban stations are included, it is important to note that the regional model is not expected to resolve the sub-grid variability required to accurately capture urban station measurements (see Section 3). These urban observations are included primarily to support the analysis, with the understanding that perfect agreement with the model is not expected. The empirical coverage is simply the percentage of observations which falls within the prediction interval. We are interested in this as a measure to ascertain, insofar as possible, the effectiveness of the method. In theory, if the model is unbiased and all sources of uncertainty have been properly accounted for, the empirical coverage should converge to the nominal (theoretical) coverage at the specified confidence level as the number of observations increases. However, in this method we only account for two sources of uncertainty: uncertainty in the amount of emissions and the temporal profile. So we should not expect that these quantities agree. However, a comparison with observations can still provide insight to the spread of the observations in relation to the prediction intervals, as well as inform us if the constructed prediction intervals are too large.

5.8.1 Bratislava 2019.02.12 - 2019.02.18

The build-up of the $PM_{2.5}$ episode in Bratislava is generally captured well by the model, which makes it a good candidate episode to analyze as a reference. This air pollution episode is primarily driven by primary particulate matter arising from combustion in the residential heating sector, with Slovakia contributing the most starting from the afternoon of 16th February. This is well reflected by the increased contribution from local emissions to the prediction interval. In this case, the prediction interval often spans many of the observations when the model bias is low. The empirical coverage is 80.56%, which is relatively high because of the generally low bias of the model and relatively high uncertainty in the residential heating sector which dominates the contributions to the episode. The spread of the prediction interval generally corresponds well with the spread in observations, even when the model exhibits some bias.

[9] A Gaussian error distribution is a common and often reasonable assumption, particularly in models where the total error arises from the aggregation of numerous independent random sources, consistent with the Central Limit Theorem. This assumption provides analytical convenience and aligns with many real-world modeling contexts. However, if the model inputs exhibit significant correlations or if the error sources are not independent or identically distributed, the Gaussian assumption may no longer hold, potentially leading to biased or mischaracterized prediction interval estimates. This can manifest, for example, in prediction intervals which contain negative values. In principle, this should not occur if the variance is correct, so we maintain this assumed prediction distribution for now, with the caveat that once all sources of uncertainty are accounted for, further investigations should prioritize rectifying this matter.

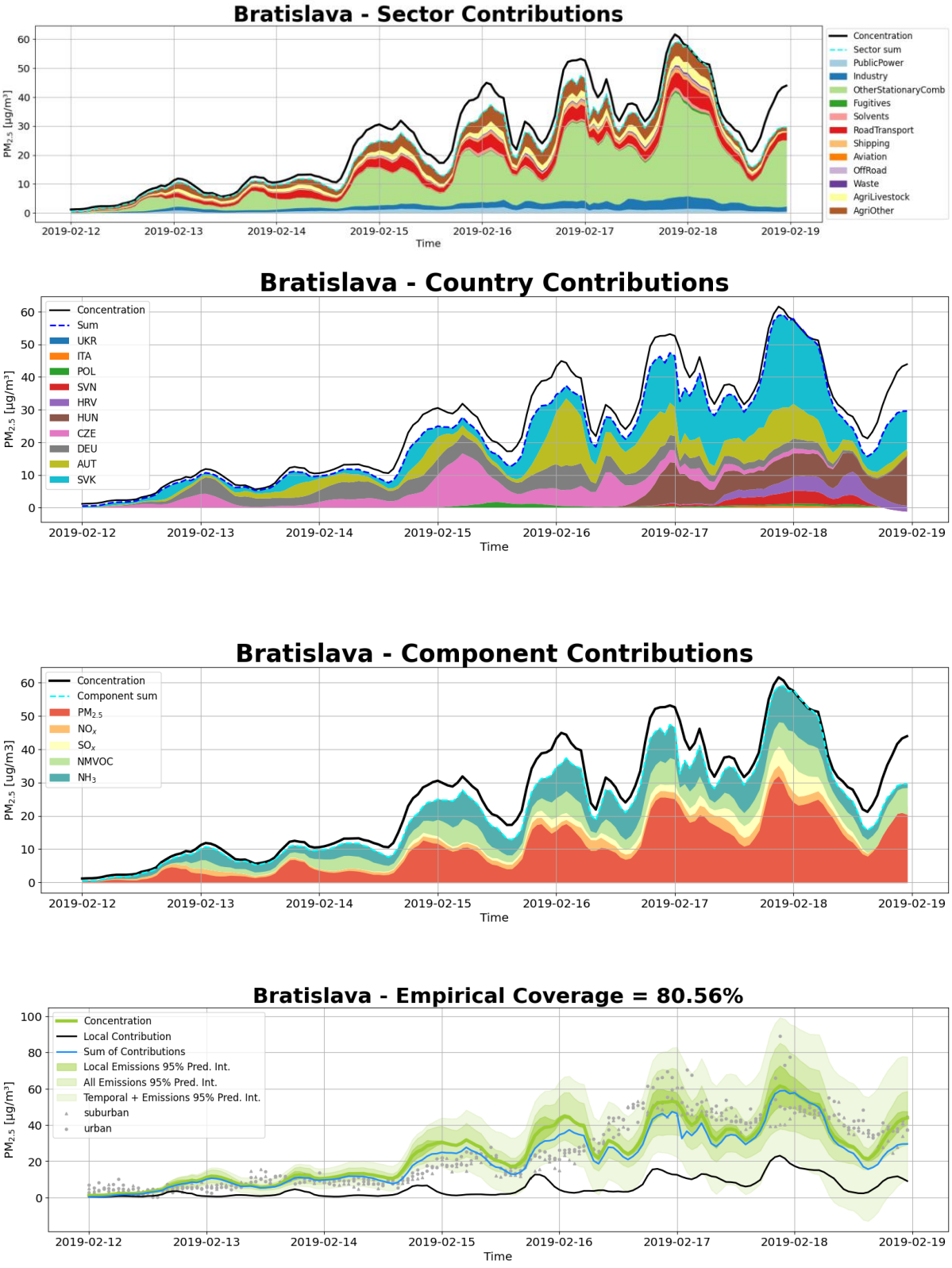


Figure 5.4: Bratislava air pollution episode partitioned into contributions by sector, country, and emitted component alongside the empirical coverage induced by the estimated prediction intervals.

5.8.2 Frankfurt 2019.02.12 - 2019.02.19

The next air pollution episode, in Frankfurt, is also primarily driven by primary particulate matter from residential heating combustion. However, significant contributions also come from aerosol precursor emissions related to agriculture, traffic, and industry. This provides valuable insight into how uncertainties across multiple sectors combine to influence the final data product. Initially, the pollution is dominated by emissions from Germany, but as the episode progresses, contributions from France and Switzerland become more prominent. The relatively even distribution of emissions across sectors suggests that national origin, rather than sector, plays a more defining role. While local emissions generally remain relatively stable throughout the event, showing typical diurnal variability, they are slightly smaller on the afternoon of the 16th when the pollution from Switzerland and Austria arrives. The empirical coverage is 79.48%, which is relatively high, but suffers because of the high bias of the model early in the episode, despite this time the spread of the observations matches the prediction interval.

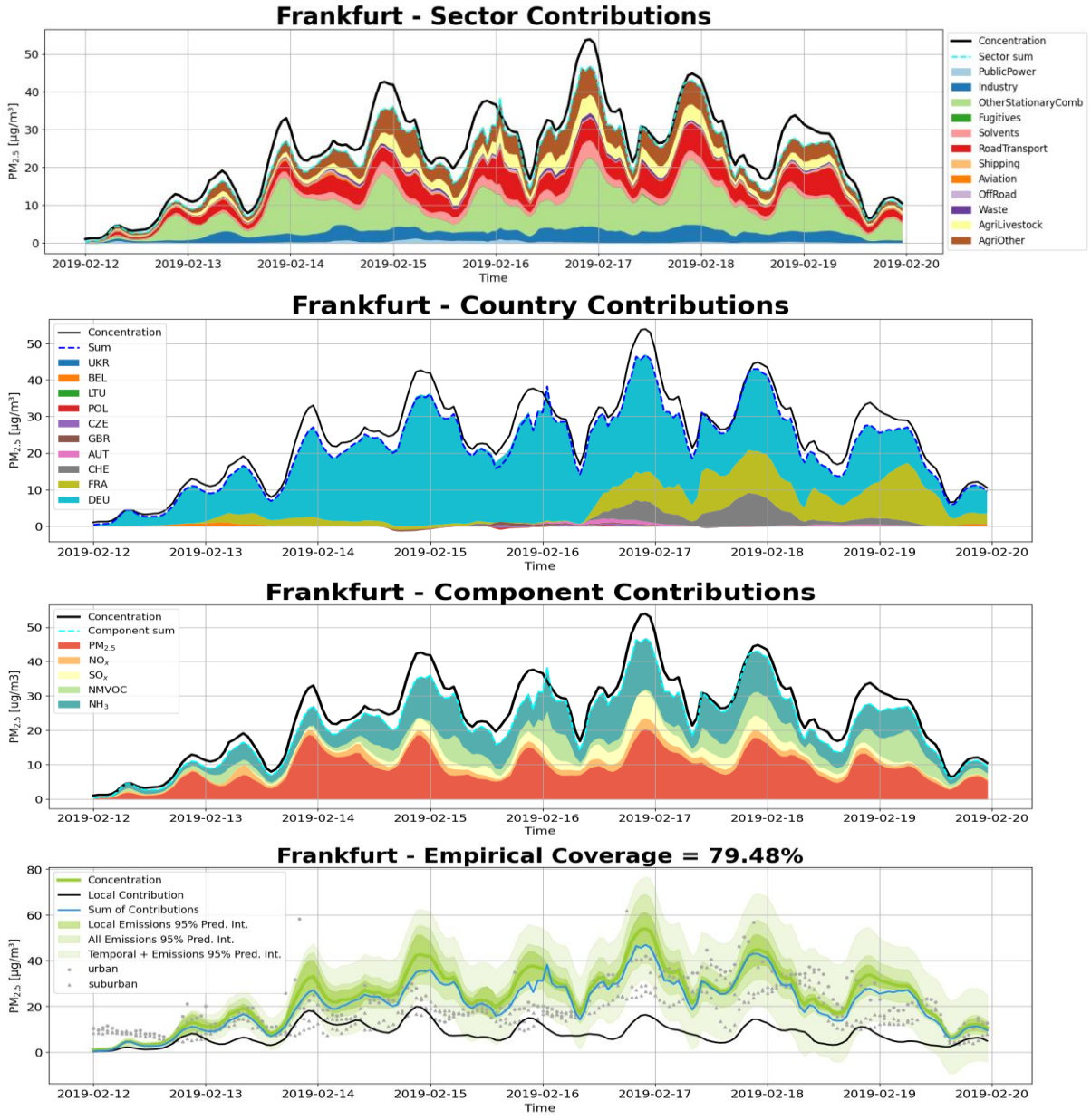
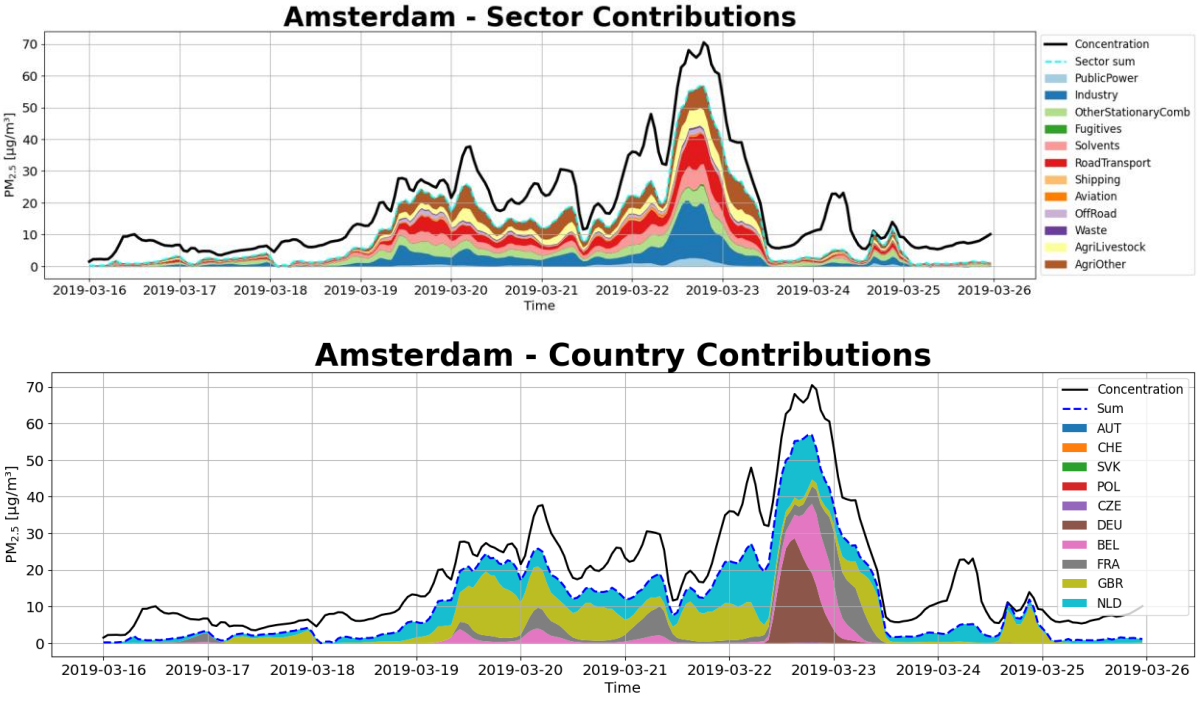


Figure 5.5: Frankfurt air pollution episode broken down by sector, country, and emitted components alongside the empirical coverage induced by the estimated prediction intervals.

5.8.3 Amsterdam 2019.03.16 - 2019.03.25

This episode in Amsterdam is primarily driven by gaseous aerosol precursors (NO_x, NH_3 and NMVOC), offering a valuable perspective on the method's performance in cases where the dominant influences are not from primary emissions. It is marked by a transition from natural to anthropogenic emissions and then back again (based CAMS Policy Service data). The time series begins with clean marine air carrying sea salt into the city. During the periods when natural sources are dominating, the prediction interval shrinks, as we do not account for uncertainties from natural emissions. Emissions arriving from the United Kingdom begin to dominate on the 19th, resulting in a wider prediction interval. As wind directions shift to the South and Southeast, air masses laden with high levels of anthropogenic pollutants, mainly ammonium nitrate and sulfate from Germany and Belgium, sweep into Amsterdam, triggering a significant pollution episode. Eventually, the winds shift once more, dispersing the pollution and restoring cleaner, more natural air conditions.

Contrasted with the previous two episodes, residential heating plays relatively little role in this episode. The peak $PM_{2.5}$ concentrations are mostly due to industrial emissions, followed by traffic, agriculture and solvents. While industry in general has a relatively high relative uncertainty, traffic and the agricultural sectors are relatively more certain. The consistent inclusion of natural sources in this episode results in prediction intervals which are small, achieving only a 43.06% empirical coverage. Some of this is due to model bias, for example on the 19th and 20th, when emission from the United Kingdom dominates. Although the bias is low during the beginning and end of the episodes, as we do not include uncertainty in the natural sources which dominate during these times, the prediction intervals are small and do not capture the spread of the observations at these times.



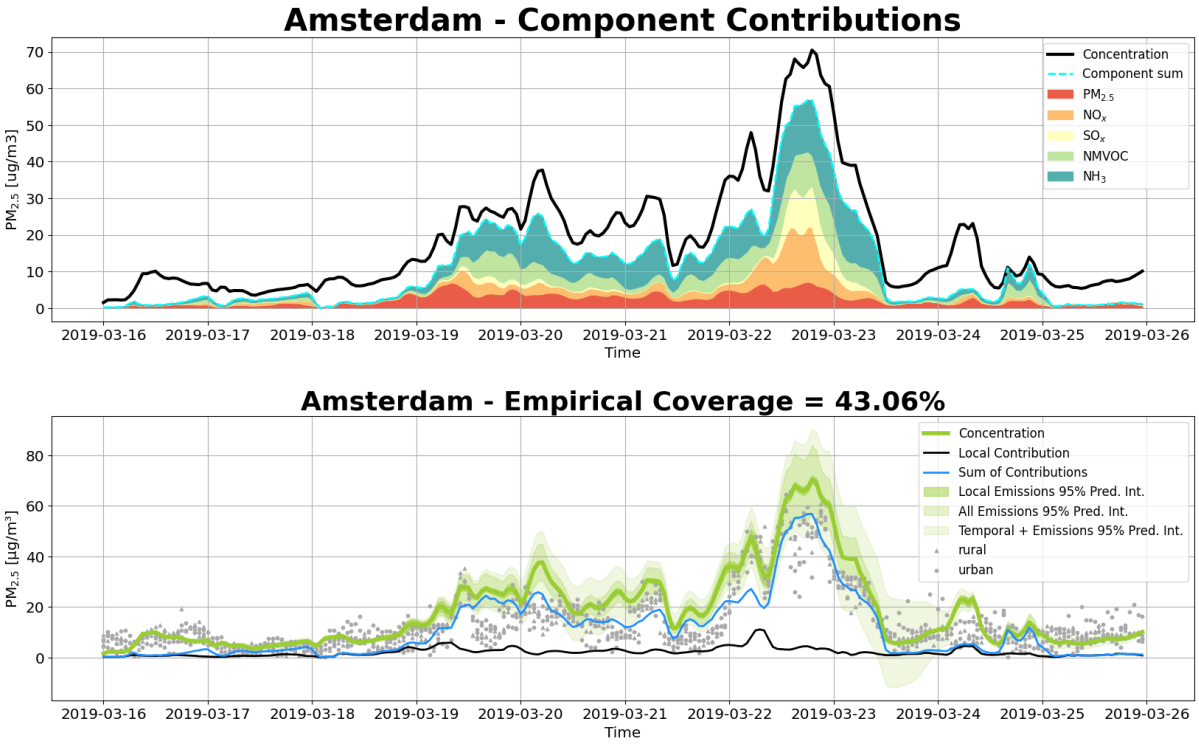


Figure 5.6: Amsterdam air pollution episode partitioned into contributions by sector, country, and emitted component alongside the empirical coverage induced by the estimated prediction intervals.

5.8.4 Warsaw 2019.10.19 - 2019.10.24

This episode in Warsaw is primarily driven by gaseous aerosol precursors (NO_x, NH_3 and NMVOC), offering another valuable perspective on the method's performance in cases where the dominant influences are not from primary emissions. There are some systemic biases in the model during the episode, which the method generally is not able to capture. Initially the episode is dominated by Polish residential heating. However, on the evening of October 22nd, contributions from the public power, industry, and traffic sectors increased. At the same time, local contributions decrease as emissions from Germany and Czechia become more significant. This shift also leads to a reduced local contribution to the prediction interval, even though the interval itself remains relatively wide due to the rather large uncertainties associated with these German and Czech emissions. Later in the episode, local residential heating emissions once again dominate, resulting in a prediction interval largely shaped by uncertainty in local emissions, despite the model tending to overestimate actual pollution levels during this time.

The empirical coverage is 75.97%. This coverage is somewhat lower in part due to the bias below, then bias above that the model has during the two main peaks of the episode. Despite this, the width of the prediction interval generally shows good alignment with the spread of observations in most cases. There are outlier observations on the 19th and the 24th from urban stations, which the model is not intended to replicate, that also contribute to lower coverage and are outside the spread of observations from suburban stations.

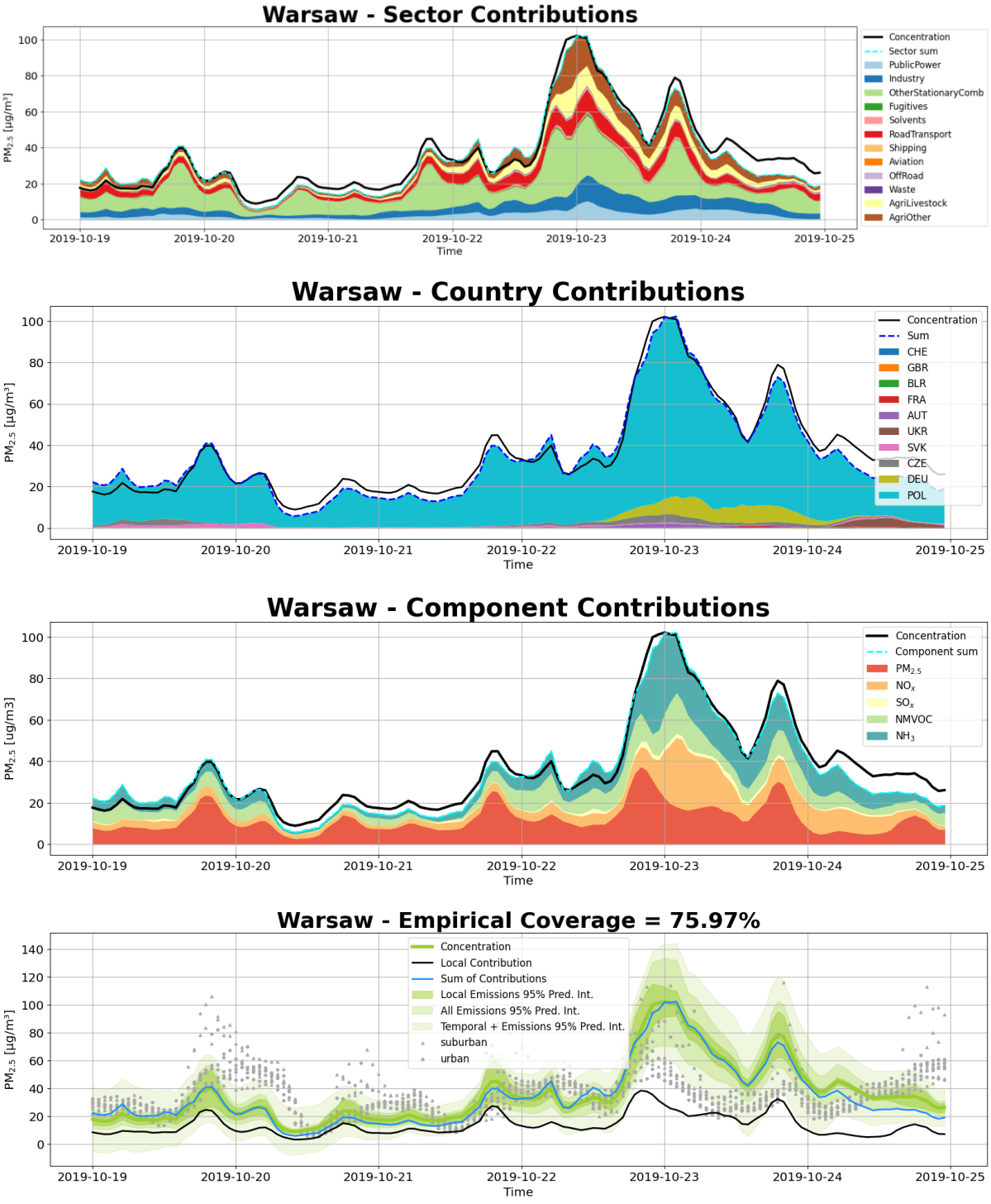


Figure 5.7: Warsaw air pollution episode partitioned into contributions by sector, country, and emitted component alongside the empirical coverage induced by the estimated prediction intervals.

5.9 Exploring Temporal Profile Aggregation Methods

In Section 5.8, each of the figures showing the empirical coverage used the same method to aggregate the temporal uncertainty: a rolling twenty-four hour window. There are other simplifications which one can consider, which can notably affect the results. Here we explore some various aggregation methods for temporal uncertainty to see the extent to which these choices matter, and what yields realistic prediction uncertainty quantification. In addition to a rolling 24 hour window, we consider the maximum, mean, and median temporal uncertainty

from the previous day as a method to aggregate these temporal uncertainties. The reason we use the day before is that it allows some time for air pollution to travel to each target city, whereas using the instantaneous temporal uncertainties from all contributing emission sources would not be realistic. In order to reliably compare the empirical coverages, each episode is run over the same period, cropping out the first day across temporal aggregation methods so that the prior day's temporal uncertainties can be leveraged.

Generally we find that the 24 hour rolling window, mean, and median yield relatively similar results, with a rather small contribution to the uncertainty derived purely from accounting for emissions uncertainty. On the other hand, aggregating the maximum uncertainty results in the most noticeable change in the total prediction interval. Without accounting for all sources of uncertainty, it remains difficult to numerically compare against observations to achieve a desired coverage.

5.9.1 Warsaw

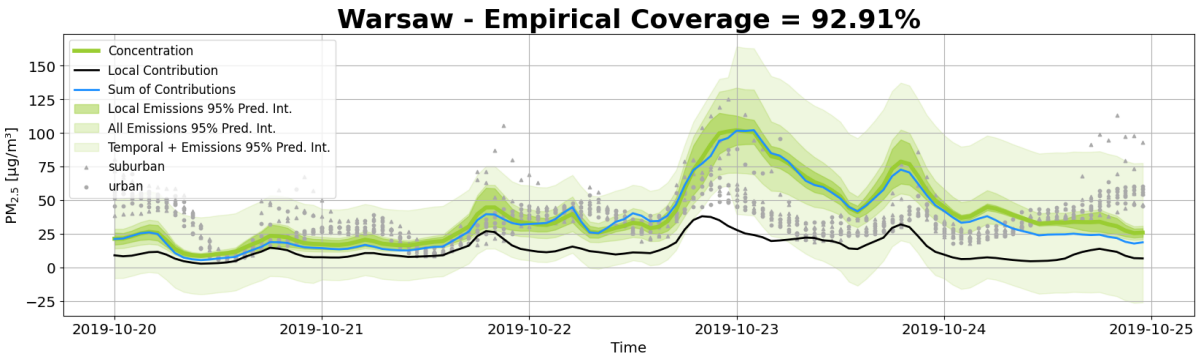


Figure 5.8: Warsaw episode: Prediction interval and associated empirical coverage achieved by aggregating the temporal profile variances using the maximum value of the previous day.

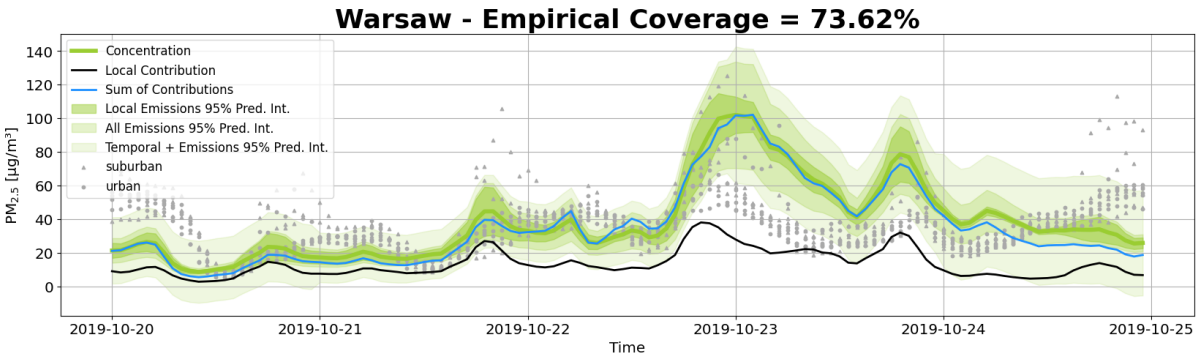


Figure 5.9: Warsaw episode: Prediction interval and associated empirical coverage achieved by aggregating the temporal profile variances using the mean value of the previous day.

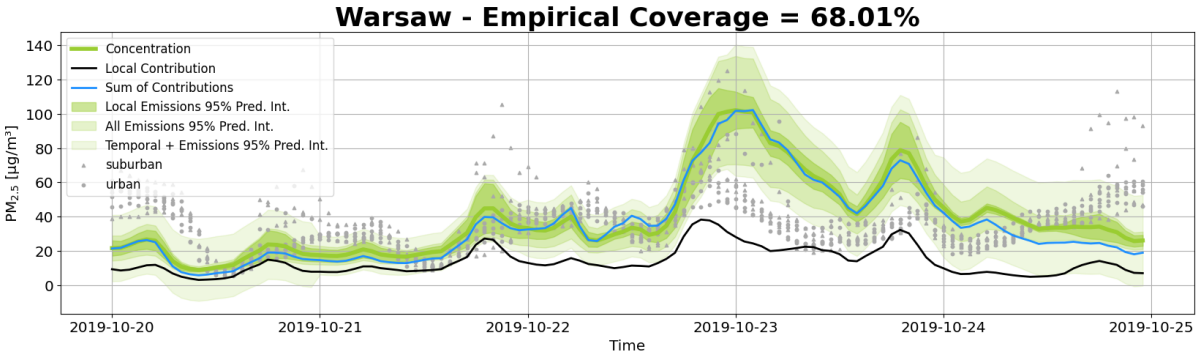


Figure 5.10: Warsaw episode: Prediction interval and associated empirical coverage achieved by aggregating the temporal profile variances using the median value of the previous day.

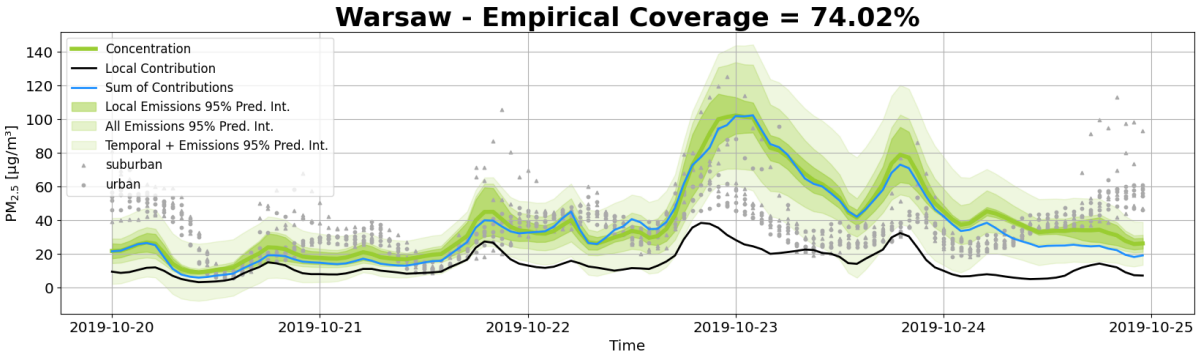


Figure 5.11: Warsaw episode: Prediction interval and associated empirical coverage achieved by aggregating the temporal profile variances using a 24 hour rolling window.

5.9.2 Bratislava

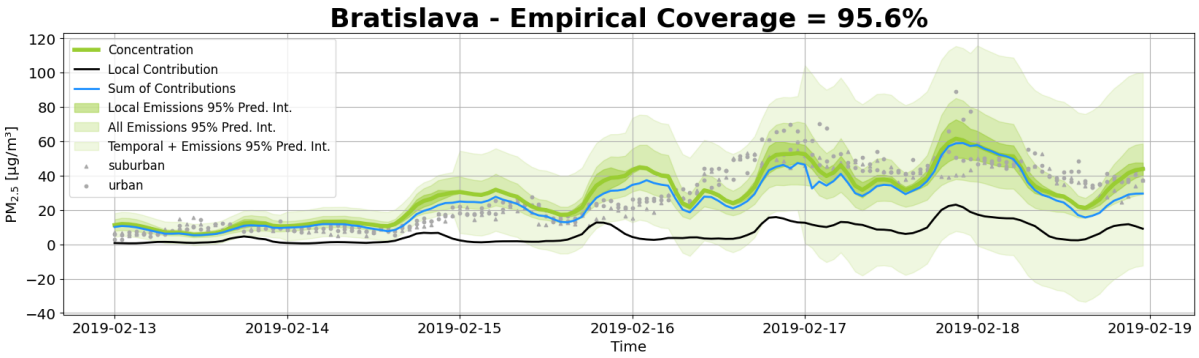


Figure 5.12: Bratislava episode: Prediction interval and associated empirical coverage achieved by aggregating the temporal profile variances using the maximum value of the previous day.

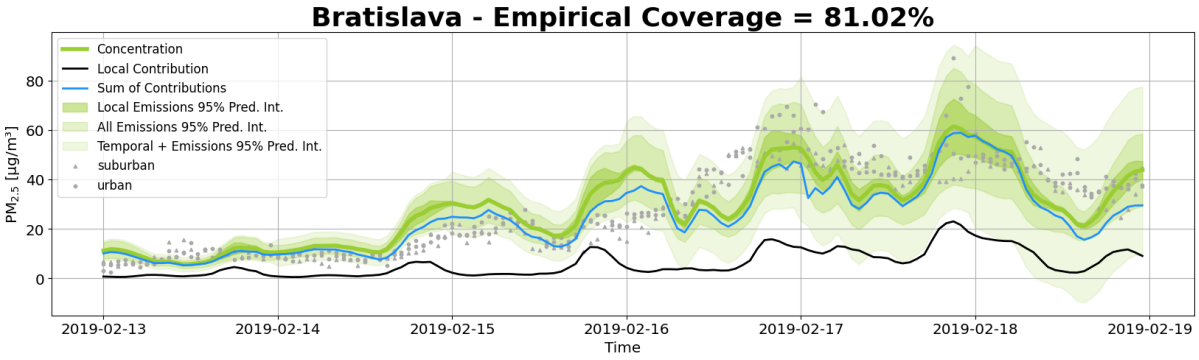


Figure 5.13: Bratislava episode: Prediction interval and associated empirical coverage achieved by aggregating the temporal profile variances using the mean value of the previous day.

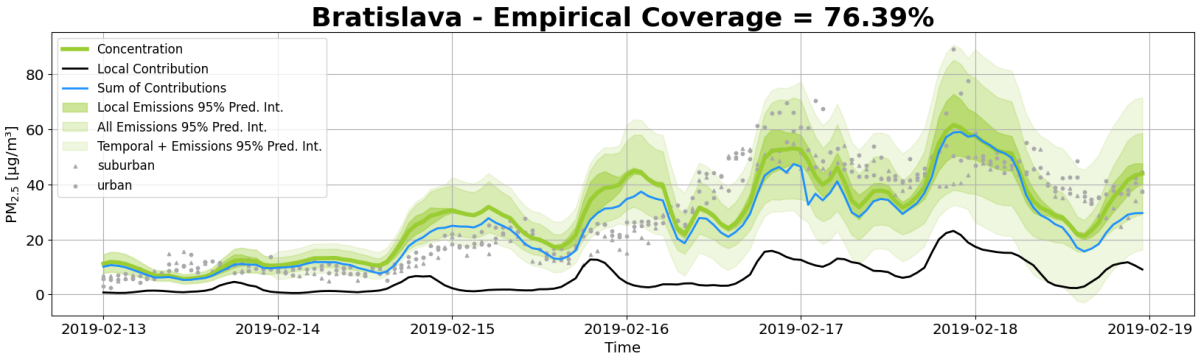


Figure 5.14: Bratislava episode: Prediction interval and associated empirical coverage achieved by aggregating the temporal profile variances using the median value of the previous day.

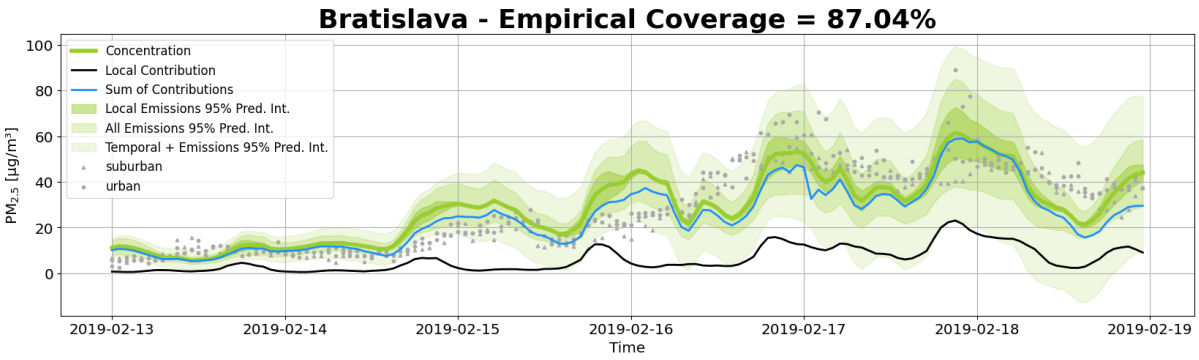


Figure 5.15: Bratislava episode: Prediction interval and associated empirical coverage achieved by aggregating the temporal profile variances using a rolling 24 hour window.

5.10 Exploring Prediction Interval Confidence Levels

It is valuable to assess empirical coverage across multiple confidence levels in order to understand how closely the observed coverage aligns with the theoretical (nominal) coverage. This helps evaluate the robustness of the predictive uncertainty estimates, beyond any single confidence threshold. Rather than relying solely on one level, typically 95%, examining alternative levels can reveal whether the method consistently captures the underlying predictive distribution. In this brief analysis, we consider 80% as an alternative level that might

be of practical relevance to users of CAMS. While 80% represents a lower confidence threshold, it is often favored in scenarios where greater tolerance for uncertainty is acceptable or where less conservative estimates are needed.

For Bratislava, the empirical coverage of the 95% prediction interval is 95.04%, matching the nominal level. This might actually indicate overestimation of the prediction variance, as several sources of uncertainty (such as those arising from natural emissions, meteorological variability, model resolution and non-linear chemical processes) are not fully accounted for in the modeling framework. The empirical coverage at the 80% prediction interval is 70.6%, reflecting a slight under-coverage. The model achieves near-nominal coverage at the 95% level (95.6%), but under-covers at the 80% level (70%). This indicates that while the extreme quantiles are well captured, the central part of the predictive distribution is too narrow, suggesting some underestimation of uncertainty for intermediate confidence levels.

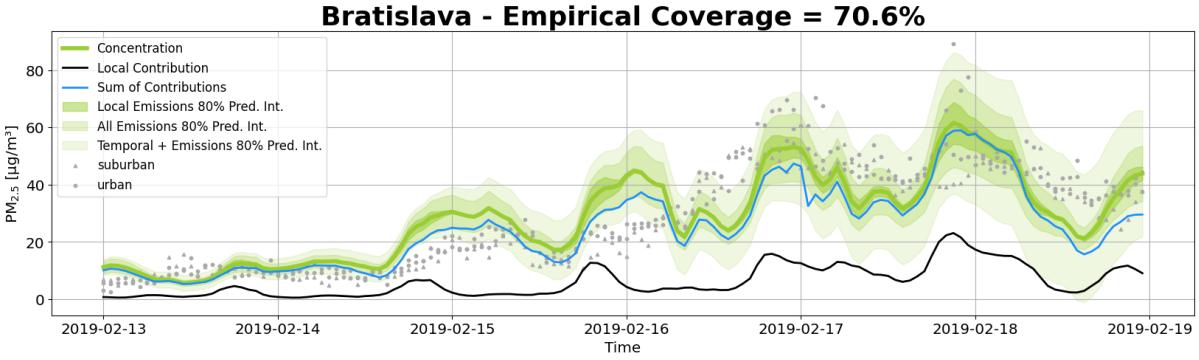


Figure 5.17: 80% Prediction interval assuming using a rolling 24 hour mean for the temporal uncertainty.

5.11 Conclusions

In this chapter, we reviewed contemporary methods for uncertainty propagation, introduced a novel analytical approach tailored for CAMS operational applications, and examined a range of correlation structures that modelers may consider. We also evaluated the method's performance across selected air pollution episodes in Europe during 2019. Overall, the approach demonstrates good performance and produces plausible uncertainty estimates, offering potential value to users of CAMS data products by quantifying the uncertainty associated with air quality forecasts.

The proposed method leverages readily available information from CAMS services, utilizes the most granular uncertainty data accessible, is grounded in mathematical rigor, and yields interpretable uncertainty estimates. Furthermore, the approach is computationally efficient, making it well-suited for rapid operational deployment. A key strength of the method is its flexibility, allowing modelers to incorporate domain-specific knowledge of realistic correlation structures. It is also extendable in that additional sources of uncertainty can be added on, allowing for refinement of the prediction variance estimates.

Despite these methodological choices, there are important limitations to keep in mind. Most critically, the underlying emissions uncertainties themselves are highly uncertain, often due to limited data, expert judgment, or assumptions in the original inventories. This adds a layer of ambiguity to any propagated uncertainty estimates. Second, the degree of correlation between emissions could significantly influence the results, but in practice, the true correlation structure is poorly understood and rarely quantified. While the correlation structure is not known, the predictive distribution can also be challenging to specify. Throughout we have used an assumed Gaussian distribution with the caveat that this can occasionally yield non-physical prediction intervals containing negative values. Future work could strive to correct this once additional sources of uncertainty have been identified and accounted for, by, for example

imposing a positive support distribution, using truncated or censored models, and considering nonparametrics and quantile methods. Lastly, our approach assumes that the relationship between emissions and modeled outputs is approximately linear. While this simplification makes the analysis more feasible, it may not fully capture complex, nonlinear interactions present in some atmospheric processes.

5.12 References

Ksenia Aleksankina, Mathew R Heal, Anthony J Dore, Marcel Van Oijen, and Stefan Reis. Global sensitivity and uncertainty analysis of an atmospheric chemistry transport model: the frame model (version 9.15. 0) as a case study. *Geoscientific Model Development*, 11(4):1653–1664, 2018.

Ksenia Aleksankina, Stefan Reis, Massimo Vieno, and Mathew R Heal. Advanced methods for uncertainty assessment and global sensitivity analysis of an eulerian atmospheric chemistry transport model. *Atmospheric Chemistry and Physics*, 19(5):2881–2898, 2019.

Luke Conibear, Carly L Reddington, Ben J Silver, Ying Chen, Christoph Knote, Stephen R Arnold, and Dominick V Spracklen. Statistical emulation of winter ambient fine particulate matter concentrations from emission changes in china. *GeoHealth*, 5(5):e2021GH000391, 2021.

Harald Cramer. *Mathematical methods of statistics*, volume 26. Princeton university press, 1999.

Alan M Dunker, Uarporn Nopmongcol, and Greg Yarwood. Uncertainty analysis of modeled ozone changes due to anthropogenic emission reductions in eastern Texas. *Atmospheric Environment*, 268:118798, 2022.

Marc Guevara, Angie Albarracın, Aitor Val, Charles Tena, Francesca Macchia, and Oriol Jorba. D5.1 Uncertainties in CAMS emission temporal profiles. Technical report, CAMS Service Evolution, 2024.

Lin Guo, Xiaokai Yang, Zhonghua Zheng, Nicole Riemer, and Christopher W Tessum. Uncertainty quantification in reduced-order gas-phase atmospheric chemistry modeling using ensemble sindy. *Journal of Geophysical Research: Machine Learning and Computation*, 1(4):e2024JH000358, 2024.

Harry H Ku et al. Notes on the use of propagation of error formulas. *Journal of Research of the National Bureau of Standards*, 70(4), 1966.

EMEP MSC-W. Opensource v5.0 (202310), October 2023. URL https://doi.org/10.5281/zenodo.8431553.

Todd Sax and Vlad Isakov. A case study for assessing uncertainty in local-scale regulatory air quality modeling applications. *Atmospheric Environment*, 37(25):3481–3489, 2003.

David Simpson, Anna Benedictow, Haldis Berge, Robert Bergström, Lisa D Emberson, Hilde Fagerli, Chris R Flechard, Garry D Hayman, Michael Gauss, Jan Eiof Jonson, et al. The EMEP MSC-W chemical transport model–technical description. *Atmospheric Chemistry and Physics*, 12(16):7825–7865, 2012.

Ingrid Super and Jeroen Keunen. M10: Uncertainties in regional emissions: First version of the European regional emissions uncertainty products available for consortium use. Technical report, CAMS Service Evolution, 2023.

Ingrid Super, Bastiaan Jonkheid, David Mathas, and Jeroen Kuenen. D5.3 uncertainties in primary PM emissions from cams-reg at the grid cell level. Technical report, CAMS Service Evolution, 2024.

P. Wind, B. Rolstad Denby, and M. Gauss. Local fractions – a method for the calculation of local source contributions to air pollution, illustrated by examples using the EMEP MSC-W model (rv4 33). *Geoscientific Model Development*, 13(3):1623–1634, 2020. doi: 10.5194/gmd-13-1623-2020. URL https://gmd.copernicus.org/articles/13/1623/2020/.

Peter Wind and Willem van Caspel. Generalized local fractions – a method for the calculation of sensitivities to emissions from multiple sources for chemically active species, illustrated using the EMEP MSC-W model (rv5.5). Geoscientific Model Development, 18(16):5397-5411, URL https://gmd.copernicus.org/articles/18/5397/2025/

Qian Ye, Jie Li, Xiao Tang, Xueshun Chen, Lei Kong, Huansheng Chen, Wenyi Yang, Huiyun Du,Xiaole Pan, Wei Wang, et al. Uncertainties in the simulated intercontinental transport of air pollutants in the springtime from emission and meteorological inputs. *Atmospheric Environment*, 293:119431, 2023.

6 Propagation of emission uncertainties into the ACT service and potential for operational implementation

6.1 Introduction

In this chapter, we present the results related to the calculation of error propagation from emissions to pollutant concentrations using the Air Control Toolbox (ACT) metamodel (Colette et al., 2022), which is widely used in the CAMS policy support framework (https://policy.atmosphere.copernicus.eu/) alongside the EMEP (Simpson et al., 2012) and LOTOS-EUROS (Manders et al., 2016) models.

The ultimate goal is to provide a methodology that can be easily translated into an operational service to be integrated into the CAMS suite for estimating uncertainties in modeled pollutant concentrations stemming from uncertainties in emissions.

The study focuses on the pollutant $PM_{2.5}$, but the methodology could be easily extended to other pollutants such as PM_{10}, NO_2 and O_3.

As already explained in the previous chapter, the standard approach for calculating error propagation in a deterministic model involves performing Monte Carlo experiments by perturbing the input variables or parameters with known uncertainties, in order to analyze how these uncertainties propagate to the output variables. However, this method is extremely computationally expensive and may be less suitable for operational systems that are required to run continuously and provide outputs every day.

Therefore, we adopted a similar approach to the one described in the previous chapter, favoring an analytical calculation of error propagation and the use of the ACT metamodel, which acts as an emulator of the CTM CHIMERE model (Couvidat et al., 2025) with respect to emission variations. One advantage of the analytical approach is the improved ability to identify the variables or sources that contribute most to uncertainty, as well as better control over the calculations being performed.

This section presents the results of the propagation of emission uncertainties to pollutant concentrations using the Air Control Toolbox (ACT) metamodel.

6.2 Total Emission and Temporal Uncertainty propagation using ACT

This section describes the methodology used to assess pollutant concentration variations, per country and sector, arising from uncertainties in temporal emission factors and annual emissions. The method is structured in the following phases:

1. Calculation of daily emission uncertainties at an hourly resolution, by pollutant and sector.
2. Mapping of these uncertainties onto the ACT model sectors.
3. Application of ACT to estimate the resulting range of variation in $PM_{2.5}$ concentrations associated with sectoral emission uncertainties.

The first step is detailed in Section 6.2.1, while the second and third steps are covered in Section 6.2.2.

6.2.1 Error propagation on hourly emissions at daily base

Two types of uncertainties associated with anthropogenic emissions were considered: those related to the total annual emissions themselves, and those related to the temporality of these total emissions.

We first calculated the uncertainty in hourly emissions based on the uncertainty in the temporal profiles provided by BSC, which were introduced in Section 5.7 of this report and are described in greater detail in Deliverable 5.1 (Guevara et al., 2024).

Starting from the annual emissions per country, pollutant, and sector, $E\_ann_{pol,sec}$, the hourly emission for a given hour and day of the year can be calculated as:

$$E\_h_{pol,sec} = \left(\frac{F\_h_{pol,sec}}{24}\right) \times \left(\frac{F\_w_{pol,sec}}{7}\right) \times \left(\frac{F\_m_{pol,sec}}{12}\right) \times E\_ann_{pol,sec} \quad \text{(Eq. 6.1)}$$

Where $F\_h_{pol,sec}$, $F\_w_{pol,sec}$, and $F\_m_{pol,sec}$ are the unnormalized hourly, weekly, and monthly temporal profiles corresponding to that specific hour and date. The division by 24, 7, and 12 ensures normalization, such that the sum of all hourly emissions $E\_h_{pol,sec}$ over the entire year equals the annual emission $E\_ann_{pol,sec}$.

To compute the uncertainty of $E\_h_{pol,sec}$ based on the uncertainties of the temporal factors$F_h_{pol,sec}$, $F\_w_{pol,sec}$, and $F\_m_{pol,sec}$ we apply the standard rules for error propagation in a multiplicative function of the form:

$$f = aABC$$

where a, A, B and C are $E\_ann_{pol,sec}$, $F\_h_{pol,sec}$, $F\_w_{pol,sec}$, and $F\_m_{pol,sec}$, respectively. Since in this step the annual emission $E\_ann_{pol,sec}$ is not treated as uncertain, it is considered a constant multiplicative parameter. Assuming that the uncertainties of the three variables are uncorrelated, it can be shown through standard mathematical derivation (Taylor, 1997) that the relative uncertainty associated with f is approximated by:

$$\frac{\sigma}{f} = \sqrt{\left(\frac{\sigma_A}{A}\right)^2 + \left(\frac{\sigma_B}{B}\right)^2 + \left(\frac{\sigma_C}{C}\right)^2} \quad \text{Eq. 6.2}$$

WIth σ_A ,σ_B, σ_C the uncertainties associated with variables A, B and C respectively.

Returning to Eq. 6.1, and based on Eq. 6.2, the relative uncertainty of hourly emissions due to the uncertainties in the temporal profiles can be expressed, for each country, pollutant, and emission sector, as follows:

$$\frac{\sigma_{pol,sec}}{E\_h_{pol,sec}} = \sqrt{\left(\frac{\sigma\_h_{pol,sec}}{F\_h_{pol,sec}}\right)^2 + \left(\frac{\sigma\_w_{pol,sec}}{F\_w_{pol,sec}}\right)^2 + \left(\frac{\sigma\_m_{pol,sec}}{F\_m_{pol,sec}}\right)^2} \quad \text{Eq. 6.3}$$

where $\sigma_{pol,sec}$ is the uncertainty in the hourly emissions, and $\sigma\_h_{pol,sec}$, $\sigma\_w_{pol,sec}$ and $\sigma\_m_{pol,sec}$ are the uncertainties associated with the hourly, weekly, and monthly temporal profiles $F\_h_{pol,sec}$, $F\_w_{pol,sec}$, $F_m_{pol,sec}$respectively.

Varying emission uncertainties on an hourly basis is not well suited for the ACT model, which is primarily designed to handle emission scenarios that are applied and maintained

consistently over longer periods. In addition, depending on the emission sectors, the hourly variability of uncertainties can be quite large and we observed that in some cases a maximum uncertainty is associated with the hours when emissions are lowest. It was then decided, for a given pollutant and sector, to apply the same relative uncertainty over all hours of a day. The relative uncertainty associated with the hour (hmax) with the highest hourly temporal factor ($F\_h_{pol,sec}^{hmax}$) was chosen and named $r\sigma\_h_{pol,sec}^{hmax}$ in the rest of the document

$$r\sigma\_h_{pol,sec}^{hmax} = \frac{\sigma\_h_{pol,sec}^{hmax}}{F\_h_{pol,sec}^{hmax}}$$

Therefore, defining the overall relative hourly uncertainty in emissions as $\varepsilon\_h_{pol,sec}$, Eq. 6.3 becomes:

$$\varepsilon\_h_{pol,sec} = \sqrt{\left(r\sigma\_h_{pol,sec}^{hmax}\right)^2 + \left(\frac{\sigma\_w_{pol,sec}}{F\_w_{pol,sec}}\right)^2 + \left(\frac{\sigma\_m_{pol,sec}}{F\_m_{pol,sec}}\right)^2} \quad \text{Eq. 6.4}$$

It is important to emphasize that this approach is both relevant and applicable because the simulation system used by INERIS — also employed to train the ACT model — now relies on CAMS-TEMPO. As a result, total annual emissions are effectively converted into hourly emissions using the same temporal profiles presented in Guevara et al. (2024).

It should be noted that this is one possible approach, but not the only one, for deriving a representative value of hourly relative uncertainty that reflects the variability over an entire day. Alternative methods could use:

1. the relative uncertainty calculated from the daily maximum absolute uncertainty
2. the maximum relative uncertainty over the day
3. the average relative uncertainty over the day

Method 1 (maximum absolute uncertainty) would give results very close to the implemented method because it is very often the same hours that show the highest hourly temporal factor and absolute uncertainty. Method 2 (maximum relative uncertainty) would overestimate uncertainties by occasionally selecting very high values that, in reality, apply only to low emissions (i.e. low hourly temporal factor). In method 3, the results do not differ significantly from those of the chosen method. The key distinction is that the chosen method is better suited to prioritizing the values associated with the strongest temporal factors, rather than assigning equal weight to all factors as in method 3.

In a second step, the uncertainty associated with annual emissions was added to the uncertainties calculation with the assumption that these uncertainties are not correlated with those associated with the time profiles.

For the ACT model, the most appropriate uncertainties are those provided—within the scope of this project—by TNO for each country, sector, and pollutant (Super et al., 2023; 2024). These country-level uncertainties offer a sufficiently smooth spatial representation (at least as a first-order approximation) to remain consistent with the fundamental assumptions underpinning the ACT meta-model.

The starting point remains Eq. 6.1, except that now the variable $E\_ann_{pol,sec}$ is also treated as uncertain.
By repeating the same steps as before, the resulting relative uncertainty—by country, pollutant, and emission sector—that accounts for both the temporal emission profiles and the annual total emissions is given by the following formula:

$$\varepsilon\_h\_tot_{pol,sec} = \sqrt{\left(r\sigma\_h\ {}^{hmax}_{pol,sec}\right)^2 + \left(\frac{\sigma\_w\ {}_{pol,sec}}{F\_w\ {}_{pol,sec}}\right)^2 + \left(\frac{\sigma\_m\ {}_{pol,sec}}{F\_m\ {}_{pol,sec}}\right)^2 + \left(\frac{\sigma\_E\_ann_{pol,sec}}{E\_ann_{pol,sec}}\right)^2}$$

TNO uncertainties are provided as annual estimates in both absolute and relative standard deviations. However, following the recommendations of the data providers, we exclusively use the relative standard deviations, denoted as $\varepsilon\_E\_ann_{pol,sec}$. The equation used to compute the total relative uncertainty $\varepsilon\_h\_tot_{pol,sec}$ is therefore as follows:

$$\varepsilon\_h\_tot_{pol,sec} = \sqrt{\left(r\sigma\_h\ {}^{hmax}_{pol,sec}\right)^2 + \left(\frac{\sigma\_w\ {}_{pol,sec}}{F\_w\ {}_{pol,sec}}\right)^2 + \left(\frac{\sigma\_m\ {}_{pol,sec}}{F\_m\ {}_{pol,sec}}\right)^2 + \left(\varepsilon\_E\_ann_{pol,sec}\right)^2} \quad \text{Eq. 6.5}$$

6.2.2 Error propagation in ACT

Before describing how sectoral emission uncertainties were computed and incorporated into ACT, we recall the formulation of the ACT model.

The Air Control Toolbox (ACT) meta-model (Colette et al., 2022) is the first air quality meta-model developed from the CHIMERE model to explore scenarios of anthropogenic emission reductions and their impacts on air quality.
ACT focuses on emissions from six main sectors: agriculture (AGR), industry (IND), residential heating (RH), road transport (TRA), shipping (SHP) and others (OTH) – including solvents, aviation, offroad, waste. It is calibrated every day using 16 perturbed CHIMERE simulations (Couvidat et al., 2025), in which emissions from the six sectors are reduced by different percentages across the entire domain to capture system variability, along with one unperturbed reference simulation.

From a mathematical perspective, ACT is a second-order polynomial that estimates the variation in pollutant concentrations based on the relative changes in emissions with respect to a reference scenario, for each emission sector.
Originally developed to estimate concentration changes for PM_{10}, $PM_{2.5}$, O_3, and NO_2, ACT has been extended within the scope of this project to also handle major PM_{10} subspecies, as detailed in Deliverable 6.3 (Zhang et al., 2025).

The complete formulation is given by:

$$\Delta C = \alpha_{agr}\,\varepsilon_{agr} + \beta_{agr}\varepsilon_{agr}^2 + \alpha_{ind}\,\varepsilon_{ind} + \beta_{ind}\,\varepsilon_{ind}^2 + \alpha_{res}\,\varepsilon_{res} + \alpha_{tra}\,\varepsilon_{tra} + \beta_{tra}\,\varepsilon_{tra}^2 + \alpha_{shp}\,\varepsilon_{shp} + \beta_{shp}\,\varepsilon_{shp}^2 + \alpha_{oth}\,\varepsilon_{oth} + \beta_{oth}\,\varepsilon_{oth}^2 + \gamma_{agr,ind}\,\varepsilon_{agr}\cdot\varepsilon_{ind} + \gamma_{tra,agr}\,\varepsilon_{tra}\cdot\varepsilon_{agr} + \gamma_{tra,ind}\,\varepsilon_{tra}\cdot\varepsilon_{ind} + \delta\,\varepsilon_{agr}\cdot\varepsilon_{ind}\cdot\varepsilon_{res}\cdot\varepsilon_{tra}\cdot\varepsilon_{shp}\cdot\varepsilon_{oth}$$

Eq. 6.6

Where:

ΔC represents the resulting change in pollutant concentration due to emission reductions. Knowing the reference concentration C_{ref}, it is possible to compute the concentration under the emission reduction scenario as $C_{agr,ind,res,tra,shp,oth}$.

$\varepsilon_{agr}, \varepsilon_{ind}, \varepsilon_{res}, \varepsilon_{tra}, \varepsilon_{shp}, \varepsilon_{oth}$ are the percentage reductions in emissions by sector, expressed as values in the range [0,1]. A value of 0 means no reduction, while a value of 1 indicates a complete reduction.

The coefficients:

$$\alpha_{agr}, \beta_{agr}, \alpha_{ind}, \beta_{ind}, \alpha_{res}, \alpha_{tra}, \beta_{tra}, \alpha_{shp}, \beta_{shp}, \alpha_{oth,}\, \beta_{oth}, \gamma_{agr,ind}, \gamma_{tra,agr}, \gamma_{tra,ind}, \delta$$

are parameters obtained by fitting Eq. 6.6 to the perturbed CHIMERE simulations.

These parameters vary by grid cell, time step, and pollutant, and are constrained such that the ACT function passes through the origin. In other words, if there are no emission changes ($\varepsilon_{agr}, \varepsilon_{ind}, \varepsilon_{res}, \varepsilon_{tra}, \varepsilon_{shp}, \varepsilon_{oth}$= 0.), the resulting concentration change must be zero ($\Delta C = 0$).

As we can see, the ACT model does not explicitly handle information related to specific chemical species. Instead, it only accounts for the percentage variation in emissions per sector.

Moreover, the emission sectors used in ACT do not follow the GNFR classification, as previously discussed, but rather an aggregated sectoral structure specific to ACT:

- AGR: Agriculture (NFR_K AgriOther + NRF_L Livestock)
- IND: Industry (NFR_A Public Power + NFR_B Industry + NFR_D Fugitive)
- RH: Residential (NFR_C OtherStationary Combustion)
- TRA: Traffic (NFR_F1 Road Transport Exhaust gasoline + NFR_F2 Road Transport Exhaust diesel + NFR_F3 Road Transport Exhaust LPG + NFR_F4 Road Transport Non exhaust)
- SHP: Shipping (NFR_G Shipping)
- OTH: Others (NFR_E Solvents + NFR_H Aviation + NFR_I Offroad + NFR_J Waste + NFR_M Other)

Therefore, the sector- and pollutant-specific relative uncertainties calculated in the previous section must be adapted before being input into ACT.

To convert emission uncertainties that depend on both pollutant and sector into uncertainties that depend only on the sector, we associate each sectoral uncertainty with the dominant emitted species for that sector.

More specifically, for GNFR sectors:

- A, F1, F2, F3, G, H, I → associated with NO_x emission uncertainties

- B → associated with SO_x emission uncertainties
- C → associated with $PM_{2.5}$ emission uncertainties
- D, E, J, F4 → associated with NMVOC emission uncertainties
- L, K → associated with NH_3 emission uncertainties

ACT sectors are, in most cases (except for Shipping and Residential), aggregations of two or more GNFR sectors. So, to each ACT sector, we assigned an uncertainty equal to the maximum of the uncertainties associated with the contributing GNFR sectors.

For example, for the agricultural sector:

$$\varepsilon_{agr} = MAX\left(\varepsilon_{GNFR_K}, \varepsilon_{GNFR_L}\right)$$

To calculate the perturbation in pollutant concentrations due to emission uncertainties, the relative errors—calculated as described above—are injected into the ACT model trained for that specific pollutant.

The hourly concentration variation per pollutant is therefore expressed as:

$$delta\_h(conc_{poll}) = \pm ACT\left(\varepsilon_{agr}, \varepsilon_{ind}, \varepsilon_{res}, \varepsilon_{tra}, \varepsilon_{shp}, \varepsilon_{oth}\right) \quad \text{Eq. 6.7}$$

In this case, ACT trained for $PM_{2.5}$ is used to propagate the uncertainties and estimate their impact on this specific pollutant.

6.3 General analysis of the results: from uncertainty on emissions to uncertainty on concentrations

The method described in 6.2 is applied to assess the impact of emission uncertainties on modeled hourly concentrations of air pollutants across various European cities.

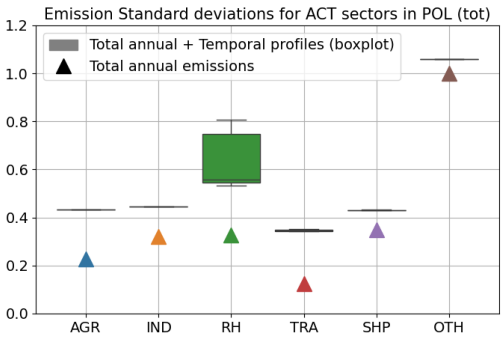
Specifically, the reference $PM_{2.5}$ concentration over each city is calculated. Then, using Equation 6.7, we estimate the resulting concentration ranges by considering either:

- the uncertainty associated solely with annual emissions, or
- the full uncertainty, including uncertainties on both annual emissions and temporal emission factors

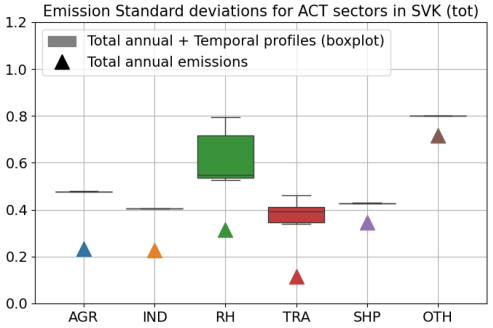
6.3.1 Emissions uncertainties as used by ACT

We first present the emission uncertainties as they are incorporated into the ACT model, already aggregated by ACT sectors. As explained in section 6.2.2, for each sector, the uncertainty associated with the most representative pollutant of the sector is selected, and if the ACT sector is made up of several GNFR sectors, the maximum uncertainty between these sectors is retained. For example, for sector RH, i.e. residential heating (also called residential sector), $PM_{2.5}$ emission uncertainties from the GNFR sector C is retained.

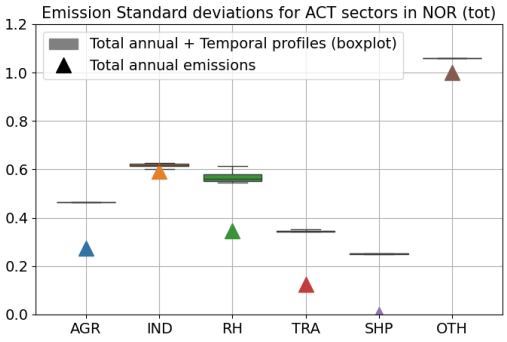
Warsaw (Poland)



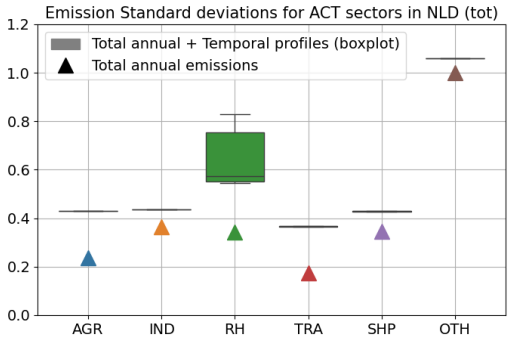
Bratislava (Slovakia)



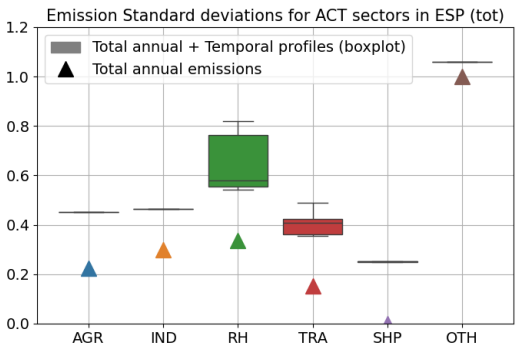
Oslo (Norway)



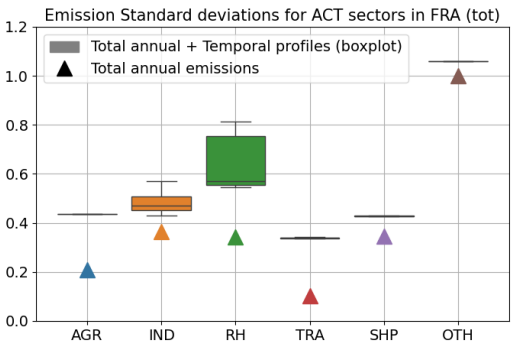
Amsterdam (Netherland)



Madrid (Spain)



Paris (France)



Milan (Italy) Brussels (Belgium)

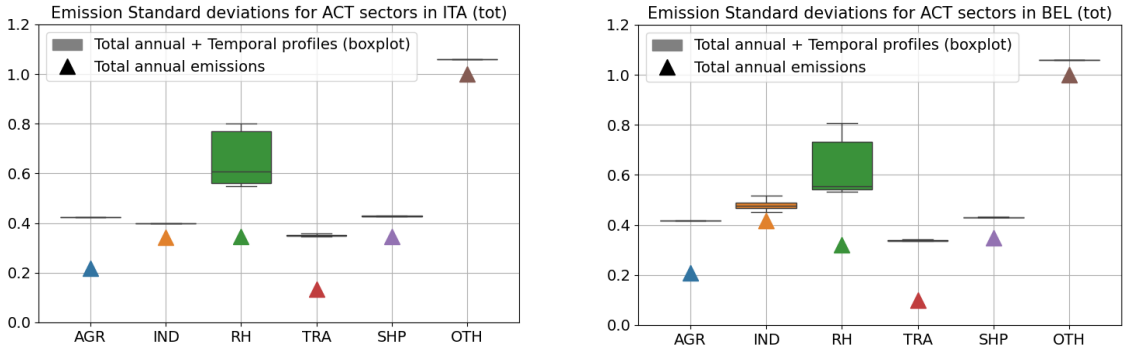


Figure 6.1 : Hourly relative uncertainties by ACT sectors (AGR, IND, RH, TRA, SHP, OTH) calculated for some selected cities over the entire year 2019. The triangles show uncertainties derived from total annual emissions uncertainties only (Super et al., 2023; 2024), while the boxplots represent the full uncertainties i.e. including both uncertainties from annual emissions and the ones associated to the temporal factors (Guevara et al., 2024).

What is striking in these figures is that the sectors with the greatest uncertainty are the RH and OTH sectors. For the latter, it is the uncertainty on the total annual emissions that induces these high values. More precisely, a detailed examination of the GNFR sectors contributing to the ACT sector (Appendix C, Figure C1), shows that it is almost always sector E (Solvents sector) which has the maximum uncertainty due to uncertainties on total NMVOC emissions. For the RH sector, the responsibility for uncertainty is shared approximately equally between that linked to total annual emissions and that linked to temporality. This sector is also the one that shows the greatest variability in emissions uncertainty over time. If we break down the temporal uncertainty (Appendix C, Figure C2), we can see that this is mainly due to the uncertainty associated with the monthly profile, which varies greatly from one month to the next. Note that for some sectors and pollutants (GNFR D, E, G, I, J etc.) the relative uncertainties from Guevara et al., (2024) do not vary over time. It is also quite notable that the uncertainty values are quite similar from one country to another, whether in terms of total or temporal uncertainty.

6.3.2 Uncertainties in concentration for cities

Figure 6.2 presents the hourly relative concentration uncertainties for each city, derived both from the total uncertainties (green box plots) and from those attributable exclusively to annual emissions (orange box plots). This is a summary figure as it represents all values for the year. However, it allows us to see if there are any significant differences between cities. The annual variability of the standard deviation of concentrations is roughly the same in these cities, with a median relative uncertainty (relative standard deviation) between 0.4 and 0.55. Some cities have less uncertainty in concentrations: Palermo (0.24), Marseille (0.27) and Athens (0.32). This is probably due to the fact that the sectors that contribute most to $PM_{2.5}$ precursor emissions are different for these Mediterranean cities, with a smaller share from the residential sector (which has high levels of uncertainty). However, it is important to note that the

uncertainty we estimate in this study is only the uncertainty of anthropogenic emissions, whereas these cities are more affected than many others by natural emissions such as, for instance, desert dust.

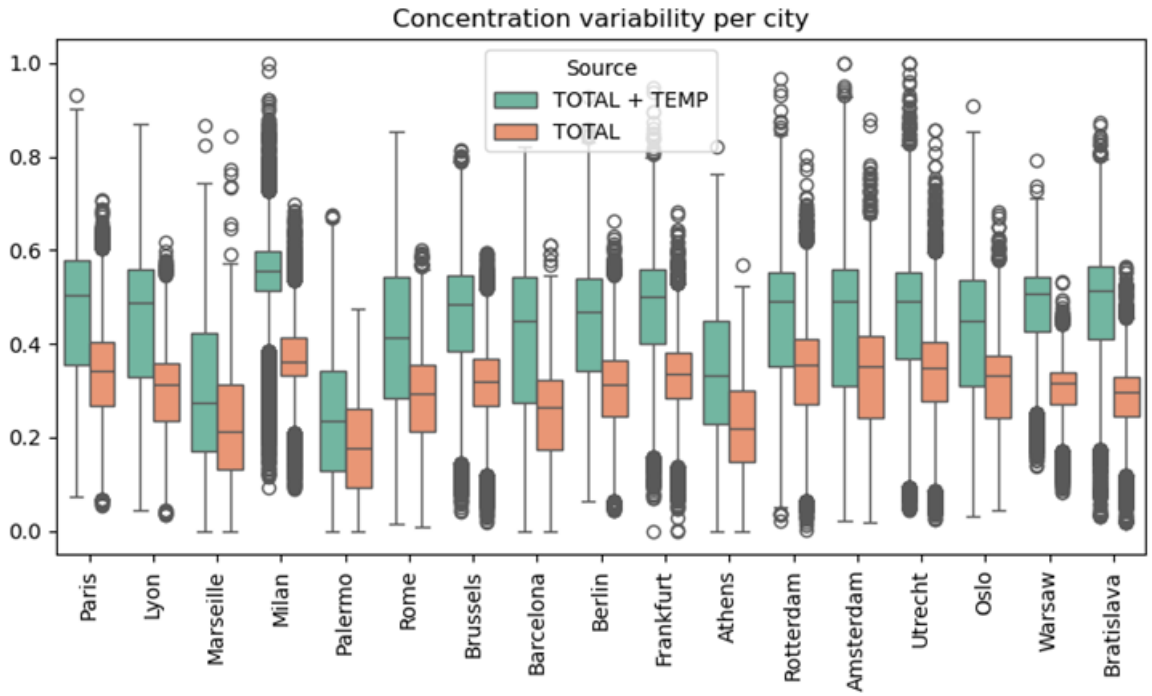


Figure 6.2 : Box plots of the relative uncertainties calculated for $PM_{2.5}$ concentrations from sectoral uncertainties on emissions for selected cities. Orange box plots represent the relative uncertainties associated with the uncertainties on total annual emissions only and the green box plots the concentrations associated with the full uncertainties (total annual + temporal factors). All hourly concentrations of the year 2019 are represented in the box plots.

This figure also allows for a quick comparison between the concentrations obtained with the complete uncertainties and those with the only annual emission uncertainties. Taking into account temporal uncertainty in addition to that in the annual total leads to an increase in uncertainties, with medians increasing by 25 to 85% depending on the city. Once again, the cities least affected by the inclusion of temporal uncertainty are the Mediterranean ones, for the same reasons mentioned earlier.

6.4 Case studies

We selected the same cities and air pollution episodes presented in Table 5.3. The aim is not only to present the results obtained with ACT, but also, through comparison with the outcomes in section 5, to evaluate how different modeling approaches and methodologies can lead to varying results.

We present the findings for the cities of Bratislava, Frankfurt, Amsterdam, and Warsaw focusing on $PM_{2.5}$ concentrations.

A comparison with observed data is included to evaluate where the actual measurements fall within the modeled concentration ranges. The observational data come from EEA (European Environment Agency) Air Quality monitoring stations and are the same as those used for model validation (https://policy.atmosphere.copernicus.eu/daily/model-evaluation). Only background monitoring stations were considered for this analysis.

In addition to quantifying the total uncertainty and that associated with the annual emission uncertainties, the ACT model is also used to isolate the contribution of each of the six ACT sectors (AGR, IND, RH, TRA, SHP, AND OTH) to the overall concentration uncertainty. By applying Equation 6.7 individually to each ε_{sec} (defined as the relative uncertainty associated with the emissions of a given sector), the sector-specific impact on modeled $PM_{2.5}$ concentrations was computed. These individual contributions are then normalized to the total concentration uncertainty, thereby providing a sectoral breakdown of the propagated emission uncertainty.

The following sections present the results for $PM_{2.5}$ concentrations across the cities of Bratislava, Frankfurt, Amsterdam, and Warsaw over the different pollution episodes.

6.3.3 Bratislava: 2019.02.12 - 2019.02.18

The pollution episode that occurred in Bratislava is also well captured by the CHIMERE model, which accurately simulates the increase in $PM_{2.5}$ concentrations between February 15 and 19 (Figure 6.3). Although observed values are significantly higher than the modeled concentrations (dark green line), they remain quite within the variability range when emission uncertainties are taken into account, with an empirical coverage of 71% when total emission errors are taken into account. It is also evident that a big contribution is given by the uncertainty originating from the temporal disaggregation of emissions (Figure 6.3 and Table 6.1). Regarding the sectoral breakdown of uncertainty (Figure 6.4), the residential sector contributes the most, with an average share of 55–57% of the total variability (column 5, Table 6.1). This is consistent with the nature of the episode, which is primarily driven by residential emissions, as also highlighted in Section 5.8.1

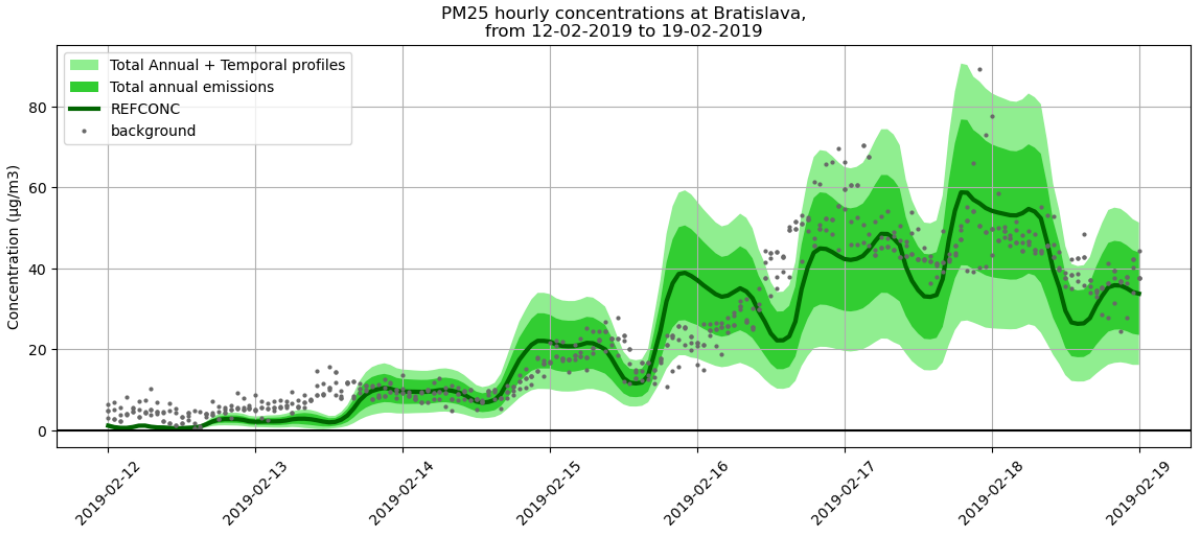


Figure 6.3: Hourly $PM_{2.5}$ concentrations during the pollution episode in Bratislava. The dark green line represents the concentrations modeled by CHIMERE. The green shaded area indicates the variability range due to uncertainties in national annual emission, while the light green shaded area represents the total uncertainty range (including both temporal factors and national total emission uncertainties). Grey dots correspond to observations from EEA background monitoring stations in Bratislava.

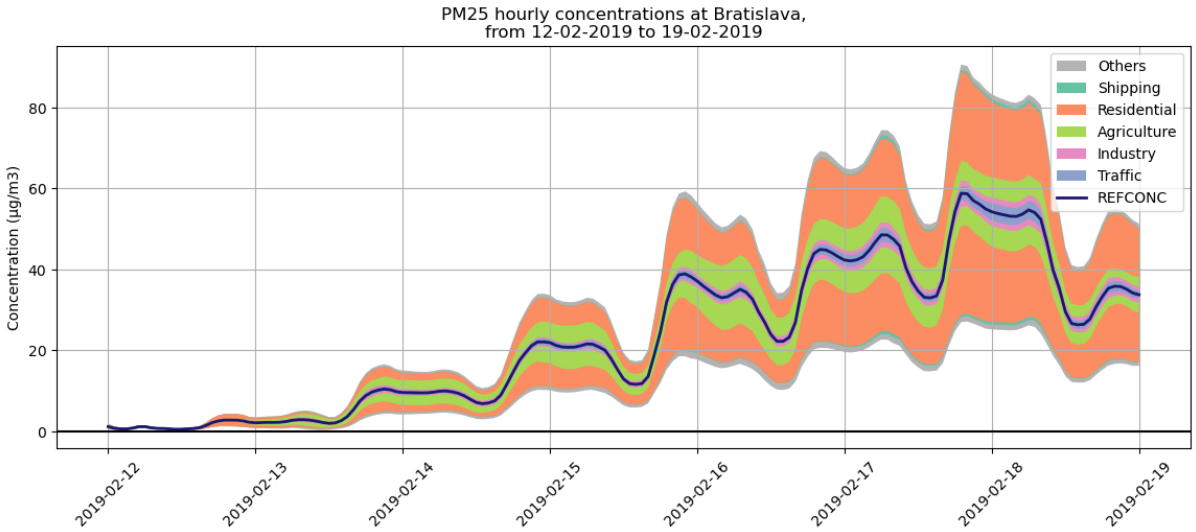


Figure 6.4 : Sectoral breakdown of the propagated emission uncertainty during the Bratislava episode. The dark green line represents the $PM_{2.5}$ concentrations modeled by CHIMERE. The colored areas indicate the contribution of emission uncertainties from different ACT sectors to the variability in $PM_{2.5}$ concentrations. Specifically, light violet represents traffic, pink represents industry, green represents agriculture, orange represents the residential sector, light blue represents shipping, and grey corresponds to other anthropogenic GNFR sectors not included in the previous categories (solvents, aviation, off-road, waste, others).

Table 6.1 : Summary table of the Bratislava episode with values averaged over the entire period. The first row refers to the results when considering the total emission uncertainty, while the second row refers to the uncertainty due to national annual emission only. The averaged variables, from left to right, are: the $PM_{2.5}$ concentration from the reference simulation, the concentration variability associated with emission uncertainty, and the percentage contribution to the uncertainty from the different ACT sectors (traffic, industry, agriculture, residential sector, shipping, and others)/ Last column refers to Empirical Coverage.

Bratislava

Err type	C_{Ref}	delta(C)	TRA	IND	AGR	RH	SHP	OTH	Empirical Coverage
	µg/m³	µg/m³	%	%	%	%	%	%	%
Total	22.9	±12.3	6	5	25	55	1	7	71
Ann	22.9	±7.0	3	5	22	57	2	11	55

6.3.4 Frankfurt 2019.02.12 - 2019.02.19

In the case of Frankfurt, CHIMERE overestimates the increase in $PM_{2.5}$ between February 14 and 16, while it aligns more closely with observations on the other days (Figure 6.5). In this case as well, visualizing the concentration variations associated with emission uncertainties supports the interpretation of the model-to-observation comparison, with an empirical coverage reaching up to 46%. For Frankfurt, the sector contributing the most to the uncertainty remains the residential sector (Figure 6.6), accounting for 31–33% of the total error. However, other sectors also contribute significantly, in order of importance: Agriculture (24–30%), Others (including solvents, aviation, offroad, and waste) with a contribution of 20–30% and Traffic (3–10%) (Table 6.2).

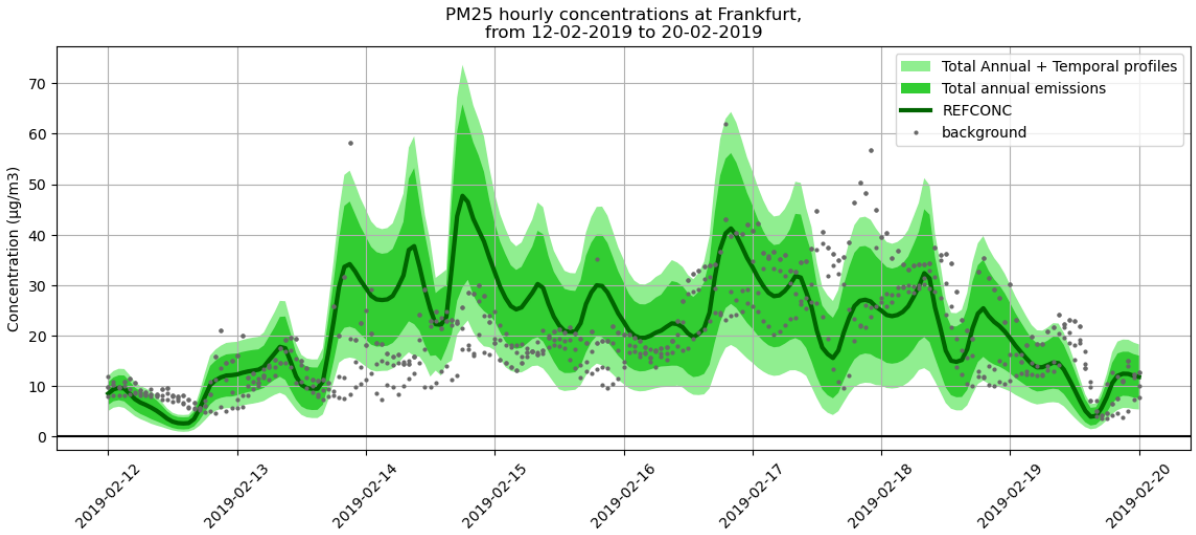


Figure 6.5 : Same as Figure 6.3, but for the pollution episode in Frankfurt.

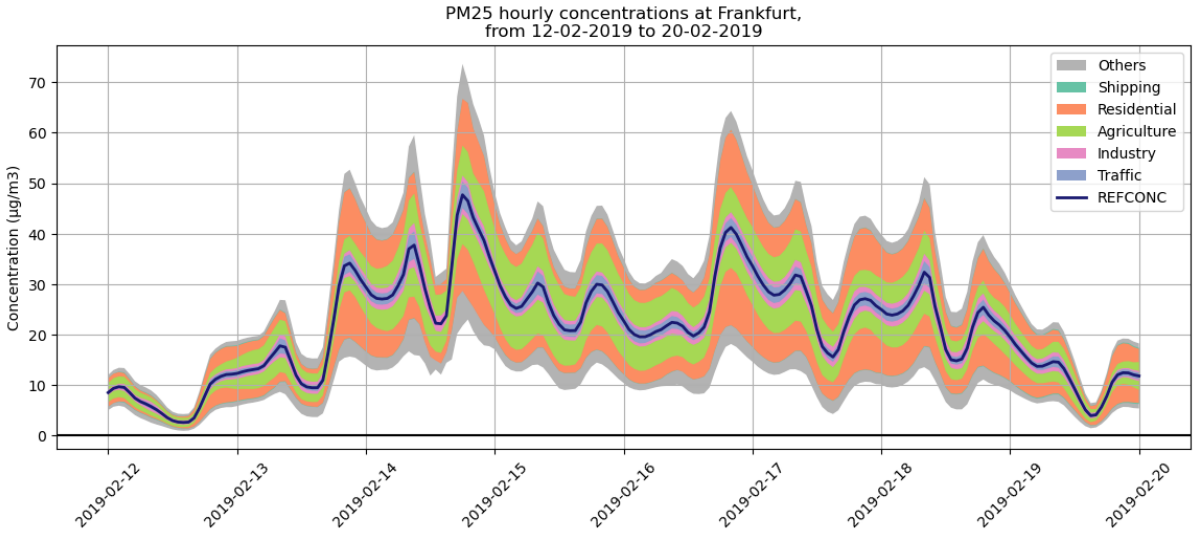


Figure 6.6 : Same as Figure 6.4, but for the pollution episode in Frankfurt.

Table 6.2 : Same as Table 6.1, but for the pollution episode in Frankfurt.

Frankfurt

Err type	C_{Ref}	delta(C)	TRA	IND	AGR	RH	SHP	OTH	Empirical Coverage
	µg/m³	µg/m³	%	%	%	%	%	%	%
Total	22.9	±11.7	10	7	30	33	0	20	46
Ann	22.9	±7.8	3	8	24	31	1	30	34

6.3.5 Amsterdam 2019.03.16 - 2019.03.25

The case of Amsterdam differs from the previous ones. As described in Section 5.8.3, this episode is not primarily driven by strong residential emissions but instead features an initial phase dominated by natural sources, followed by a second period influenced by air masses carrying anthropogenic pollutants (Figure 6.7). Since uncertainties related to natural emissions are not considered, the resulting propagated error remains low during the first days. In contrast, an increase in uncertainty is observed in the latter period. During this second phase, the main contributor to the total error is Others sector (28-39%), followed by Agriculture (23-33%) (Figure 6.8 and Table 6.3).

It is also observed that, the overall variability in $PM_{2.5}$ concentrations remains below 50%, which is lower compared to the other case studies. As will be further discussed in the following section, this is primarily due to the fact that the residential sector—associated here with the highest uncertainty—plays a less dominant role in this episode. Consequently, when emissions from the residential sector are not the main driver of concentrations, the associated propagated uncertainty tends to be reduced.

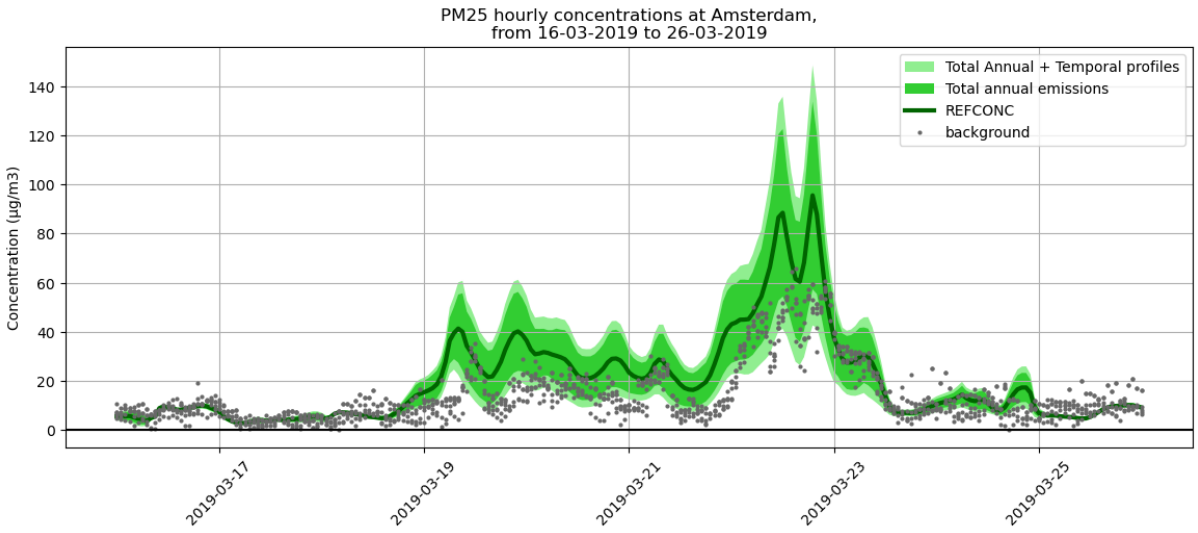


Figure 6.7 : Same as Figure 6.3, but for the pollution episode in Amsterdam.

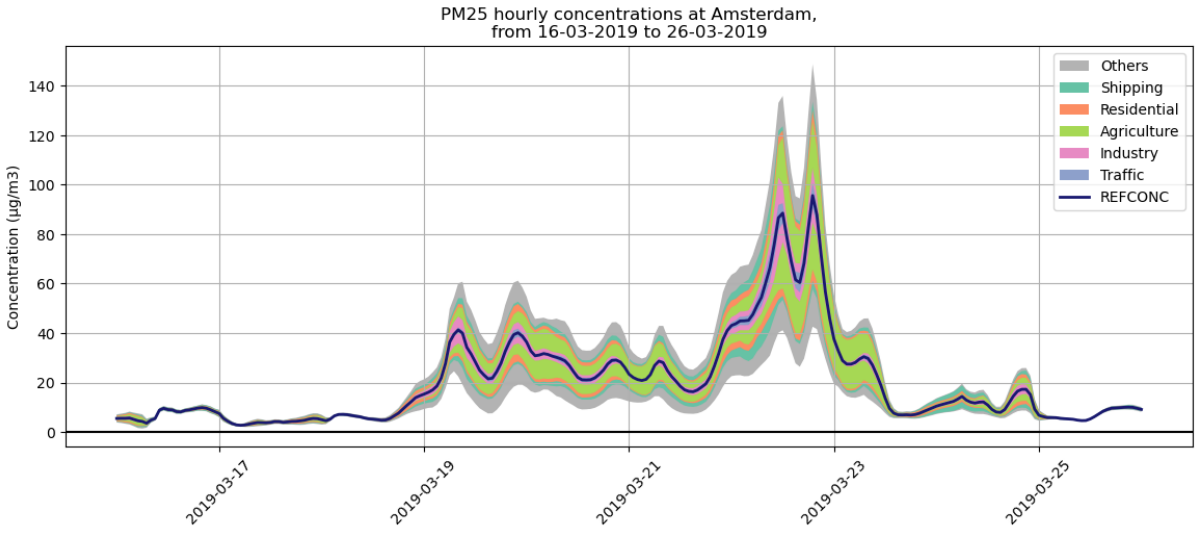


Figure 6.8 : Same as Figure 6.4, but for the pollution episode in Amsterdam.

Table 6.3 : Same as Table 6.1, but for the pollution episode in Amsterdam.

Amsterdam

Err type	C_{Ref}	delta(C)	TRA	IND	AGR	RH	SHP	OTH	Empirical Coverage
	µg/m³	µg/m³	%	%	%	%	%	%	%
Total	22.9	±9.8	7	11	33	10	11	28	49
Ann	22.9	±6.8	5	13	23	9	11	39	36

6.3.6 Warsaw 2019.10.19 - 2019.10.24

In the case of Warsaw, the model tends to overestimate the observations although it remains within the expected variability range (Figure 6.9) with an empirical coverage of 73% when total emission error is taken into account. The sector contributing the most to the propagated uncertainty is Residential, accounting for 37–38% of the total error, followed by the Agriculture, which contributes 28–32% (Figure 6.10 and Table 6.4).

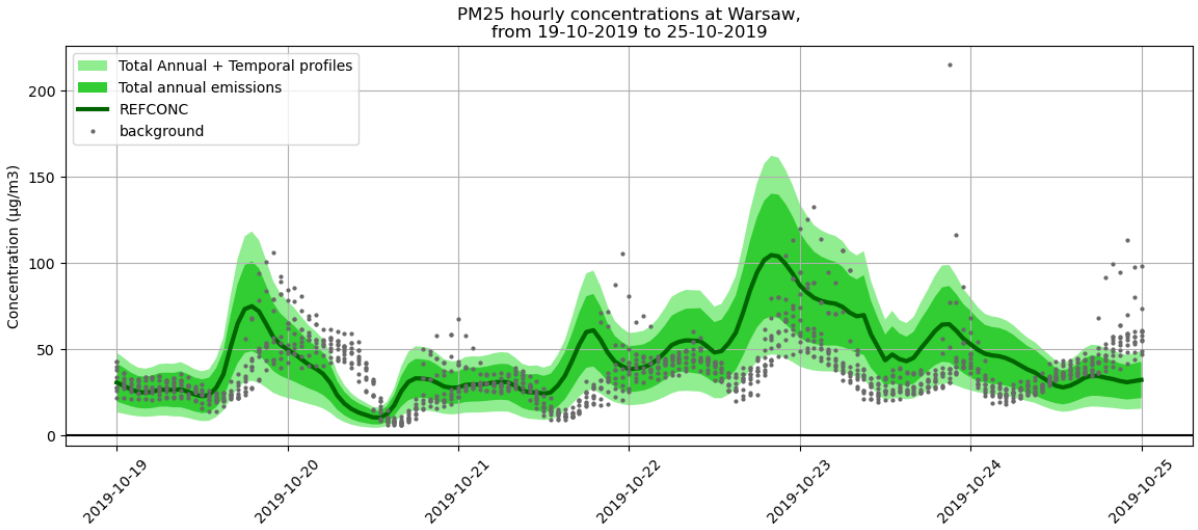


Figure 6.9 : Same as Figure 6.3, but for the pollution episode in Warsaw.

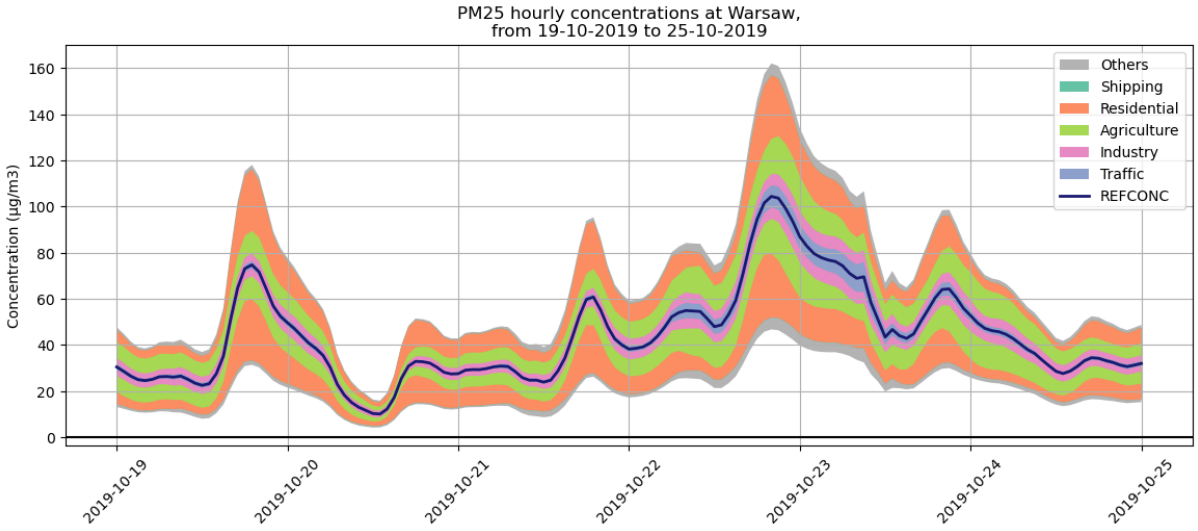


Figure 6.10: Same as Figure 6.4, but for the pollution episode in Warsaw.

Table 6.4: Same as Table 6.1, but for the pollution episode in Warsaw.

Warsaw

Err type	C_{Ref}	delta(C)	TRA	IND	AGR	RH	SHP	OTH	Empirical Coverage
	µg/m³	µg/m³	%	%	%	%	%	%	%
Total	22.9	±24.1	8	13	32	38	0	8	73
Ann	22.9	±15.4	5	15	28	37	1	14	47

6.4 Sensibility to temporalization

The issue of how to validly assign errors related to the temporal distribution of emissions does not have an obvious solution. For example, the concentrations at a specific spatial point within our domain—let's denote it as point *j*—depend on a set of emissions originating from various locations across the domain, which are not necessarily in close proximity. Consequently, the error affecting the concentration at that point may be associated with different temporal uncertainties, since the contributing emissions may not have occurred at the same time.

If local emissions are assumed to be predominant, the temporal error corresponding to the same day could be a reasonable approximation. On the other hand, if it is considered that the emissions influencing concentrations at the point are more distant, it may be more appropriate to account for the error from the previous day—or even earlier.

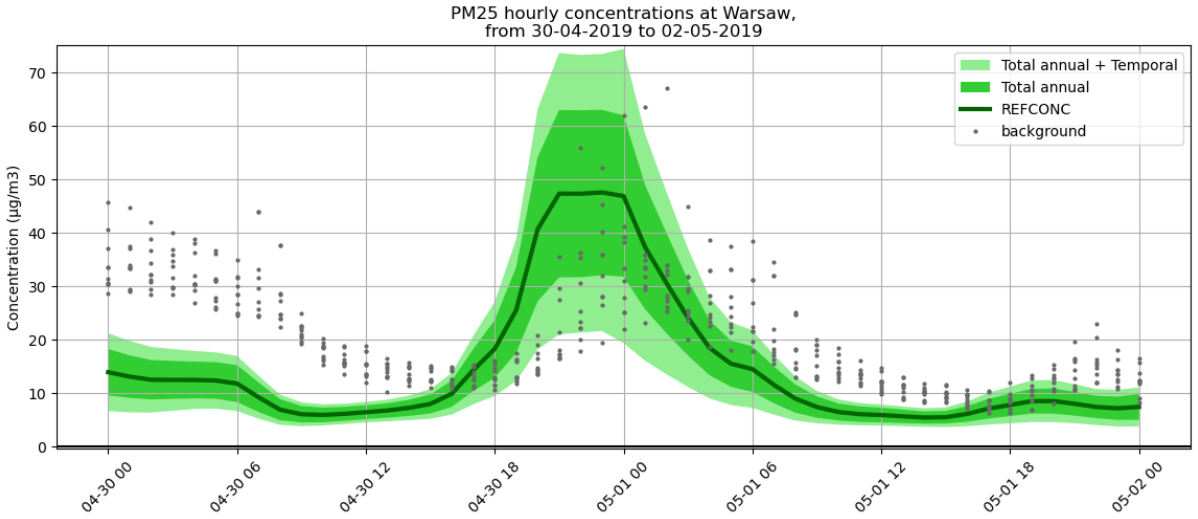
Since we are not able to modulate the temporal error of emissions based on the time it takes for them to reach point *j*, and can only apply the same temporal offset to all emissions, we decided to conduct a sensitivity test. Two scenarios were considered: (1) the emission errors from day *x* were propagated to the concentrations on the same day: *"Same day scenario"* (*x*), and (2) the emission errors from day *x–1* were propagated to the concentrations on day *x: "Day before scenario"*. The resulting $PM_{2.5}$ concentration errors in the two cases were then compared over different cities: Paris, Lyon, Marseille, Milan, Palermo, Rome, Brussels, Barcelona, Berlin, Frankfurt, Athens, Rotterdam, Amsterdam, Utrecht, Oslo, Warsaw, Bratislava.

The results show that the differences between the relative $PM_{2.5}$ concentration errors in the two scenarios are very small. Considering the cities observed throughout the entire year 2019, 75% of the hourly concentrations show a difference of only 0.04% with respect to the total $PM_{2.5}$ concentration, with a maximum difference of 4.6% observed in Warsaw on 1 May.

To better illustrate the differences between the two scenarios, Figure 6.11 presents the case of Warsaw. The top plot corresponds to the *Same Day scenario*, while the bottom plot corresponds to the *Day Before scenario*. As can be seen, the differences are barely noticeable, with a slight deviation around midnight on 1 May.

The reason why the differences are not large at all lies in the fact that, although the errors on temporal factors can be quite significant (as we have already shown), their day-to-day variability is not especially high.

Same Day scenario



Day Before scenario

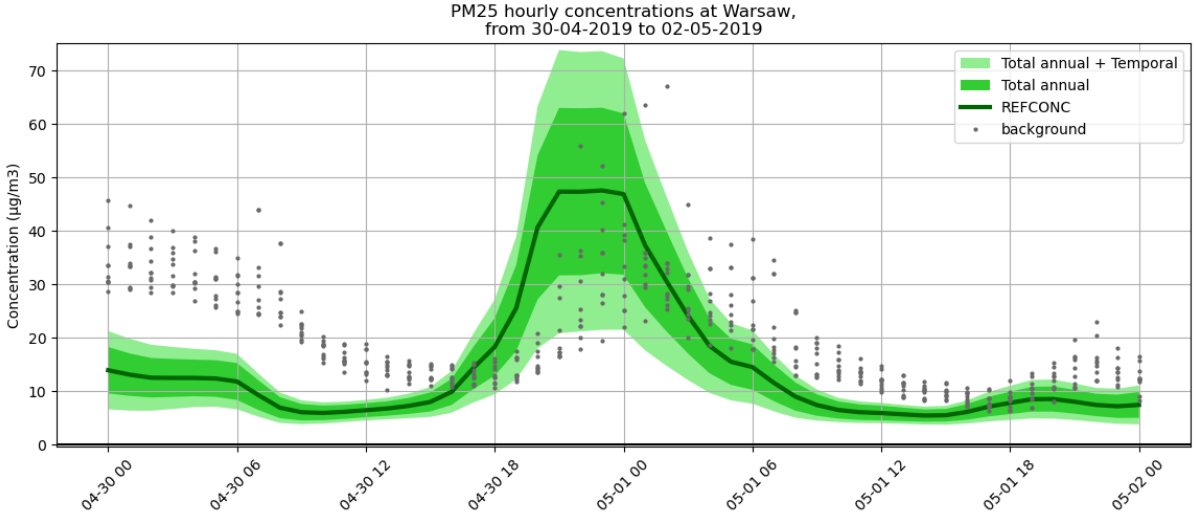


Figure 6.11 : Hourly $PM_{2.5}$ concentrations in Warsaw from 30 April to 2 May 2019 under the Same Day scenario (top panel) and the Day Before scenario (bottom panel). The dark green line represents the concentrations modeled by CHIMERE. The green shaded area indicates the variability range due to uncertainties in annual emission, while the light green shaded area represents the total uncertainty range (including both temporal factors and national annual emission uncertainties). Grey dots correspond to observations from EEA background monitoring stations in Warsaw.

6.5 Conclusion

This section presented a methodology for propagating emission uncertainties to $PM_{2.5}$ concentrations, accounting for both uncertainties in temporal allocation factors and total annual emissions. The approach leverages the ACT metamodel, starting from an initial estimate of hourly emission errors derived through analytical error propagation from temporal factors (Guevara et al., 2024), as well as from country-, sector-, and pollutant-specific annual emission uncertainties (Super et al., 2023, 2024). The current implementation provides estimates of variability in $PM_{2.5}$ concentrations over major European cities, but the methodology can be readily extended to all grid points within the modeling domain and to additional pollutants.

The method is computationally efficient, requiring minimal processing time and resources, making it particularly well suited for potential operational deployment.

Results indicate relatively large propagated errors, which are consistent with the substantial uncertainties associated with emissions—especially those related to temporal disaggregation factors. The residential sector is by far the most affected by temporal uncertainties, and consequently, the total propagated errors are highest in regions and periods where this sector dominates.

Notable differences arise when comparing this method to that described in chapter 5. These discrepancies can be attributed to several factors. While both methods adopt an analytical approach to variance estimation, they rely on different formulations of uncertainty propagation. In the present method, uncertainties are treated as independent, while the results presented in chapter 5 assume some degree of correlation between emission sources. Furthermore, the two underlying models, ACT and the Local Fraction method, differ substantially. Notably, ACT does not distinguish between local and non-local emissions, meaning that the computed

hourly emission errors are applied uniformly across the domain. This may lead to an upper-bound estimate of the concentration errors.

Further investigation is needed to fully understand the drivers of divergence between the two approaches. Nevertheless, the results underscore the fact that different methodologies for uncertainty propagation can themselves lead to substantial variability—highlighting the importance of cross-validation in air quality uncertainty assessments.

6.6 References

Colette, A., Rouïl, L., Meleux, F., Lemaire, V., and Raux, B. (2022). Air Control Toolbox (ACT_v1.0): A flexible surrogate model to explore mitigation scenarios in air quality forecasts. Geoscientific Model Development, 15(4), 1441–1465. https://doi.org/10.5194/GMD-15-1441-2022

Couvidat, F., Lugon, L., Messina, P., Sartelet, K., & Colette, A. (2025). Optimizing computation time in 3D air quality models by using aerosol superbins within a sectional size distribution approach: Application to the CHIMERE model. Journal of Aerosol Science, 187, 106572. https://doi.org/10.1016/J.JAEROSCI.2025.106572

Guevara, M., Albarracın, A., Val A., Tena C., Macchia F., and Jorba, O. D5.1 Uncertainties in CAMS emission temporal profiles. Technical report, CAMS Service Evolution, 2024.

Ingrid Super and Jeroen Keunen. M10: Uncertainties in regional emissions: First version of the European regional emissions uncertainty products available for consortium use. Technical report, CAMS Service Evolution, 2023.

Super, I., Jonkheid, B., Mathas, D., and Kuenen, J. D5.3 uncertainties in primary PM emissions from cams-reg at the grid cell level. Technical report, CAMS Service Evolution, 2024.

Ku, H. H. (October 1966). "Notes on the use of propagation of error formulas". Journal of Research of the National Bureau of Standards. 70C (4): 262. doi:10.6028/jres.070c.025. ISSN 0022-4316. Retrieved 3 October 2012.

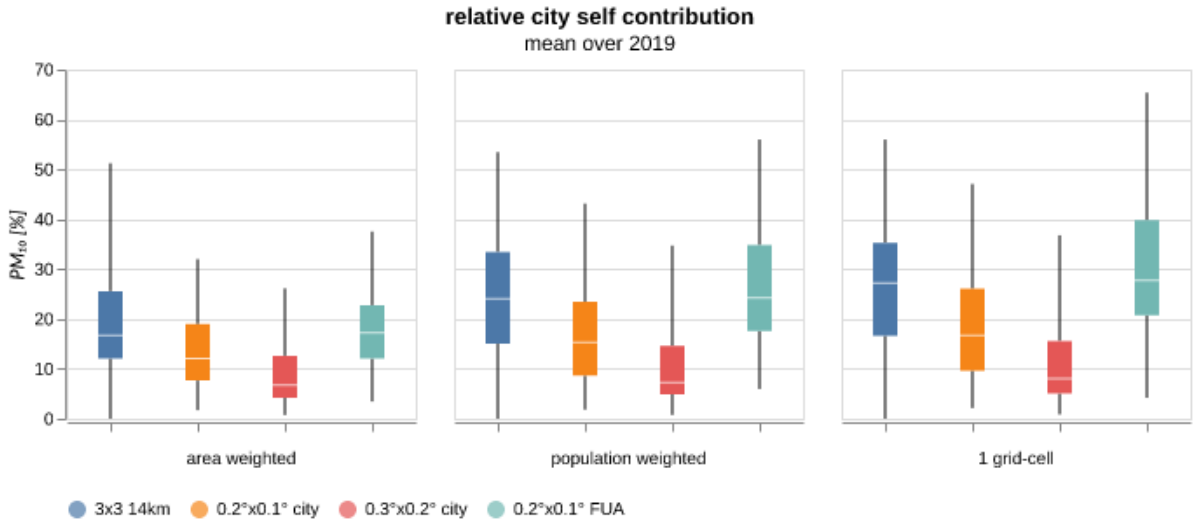
Manders, A., Builtjes, P., Curier, L., Denier van der Gon, H., Hendriks, C., Jonkers, S., Kuenen, J., Timmermans, R., Wichink Kruit, R. A., van Pul, A., Sauter, F., van der Swaluw, E., Douros, J., Eskes, H., van Meijgaard, E., van Ulft, B., Mues, A., Banzhaf, S., & Schaap, M. (2016). Curriculum Vitae of the LOTOS-EUROS chemistry transport model (v1.11). Geoscientific Model Development, submitted.

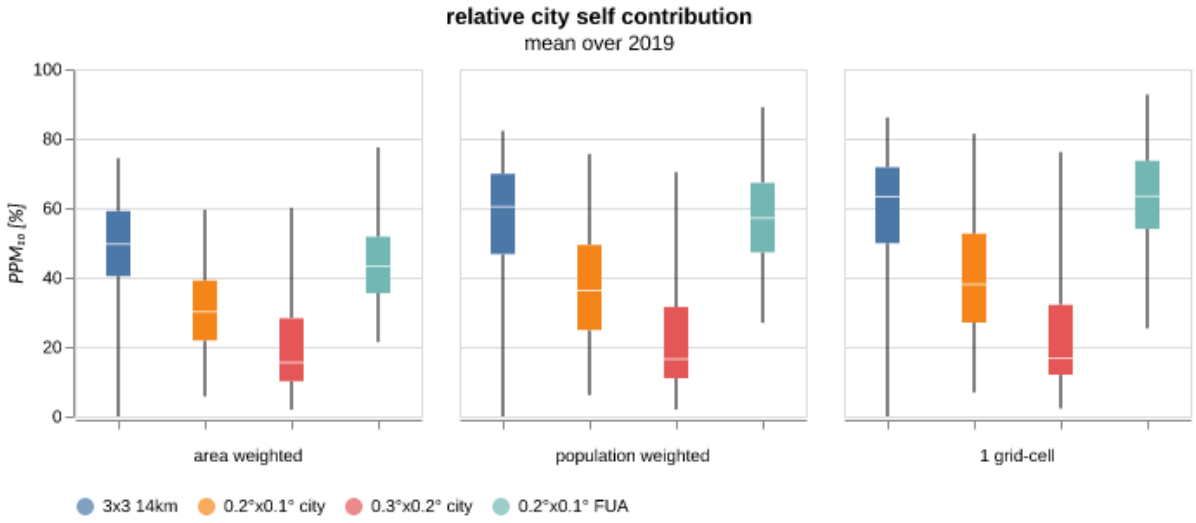
Simpson, D., Benedictow, A., Berge, H., Bergström, R., Emberson, L. D., Fagerli, H., Flechard, C. R., Hayman, G. D., Gauss, M., Jonson, J. E., Jenkin, M. E., Nyíri, A., Richter, C., Semeena, V. S., Tsyro, S., Tuovinen, J.-P., Valdebenito, Á., and Wind, P.: The EMEP MSC-W chemical transport model – technical description, Atmos. Chem. Phys., 12, 7825–7865, https://doi.org/10.5194/acp-12-7825-2012, 2012.

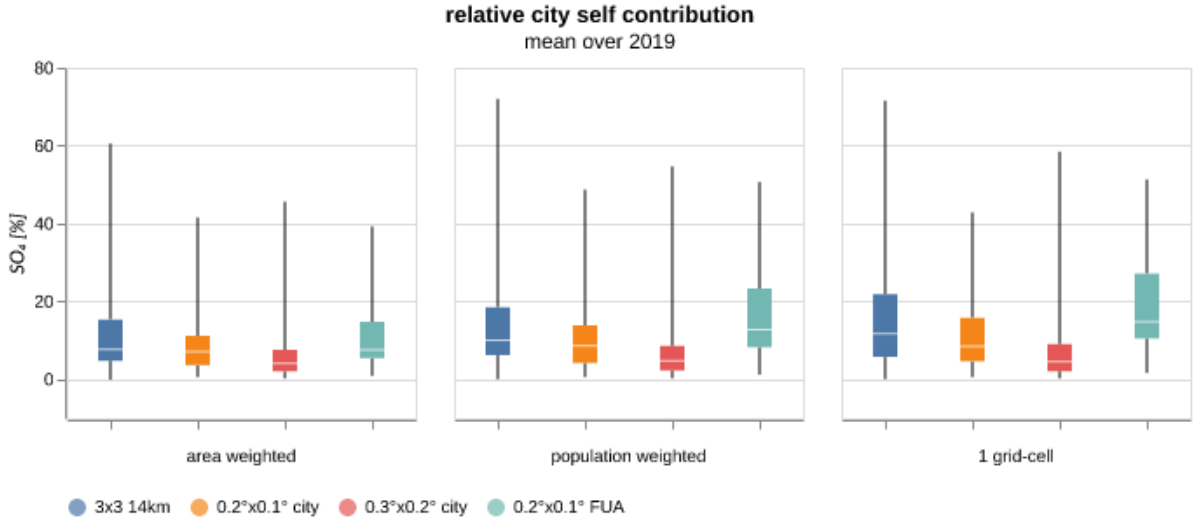
Taylor, J., Introduction to Error Analysis, The Study of Uncertainties in Physical Measurements, University Science Books, NY, 1997.

Zhang, J., Pekel, F., Timmermans, R., Tokaya, J., Fagerli, H., Valdebenito, A., Tsyro S., Colette, A., Messina, P., Raux, B., Real, E., Wind, P. D6.3 Uncertainties in Sector/Source contributions. Technical report, CAMS Service Evolution, 2025.

10. APPENDIX A







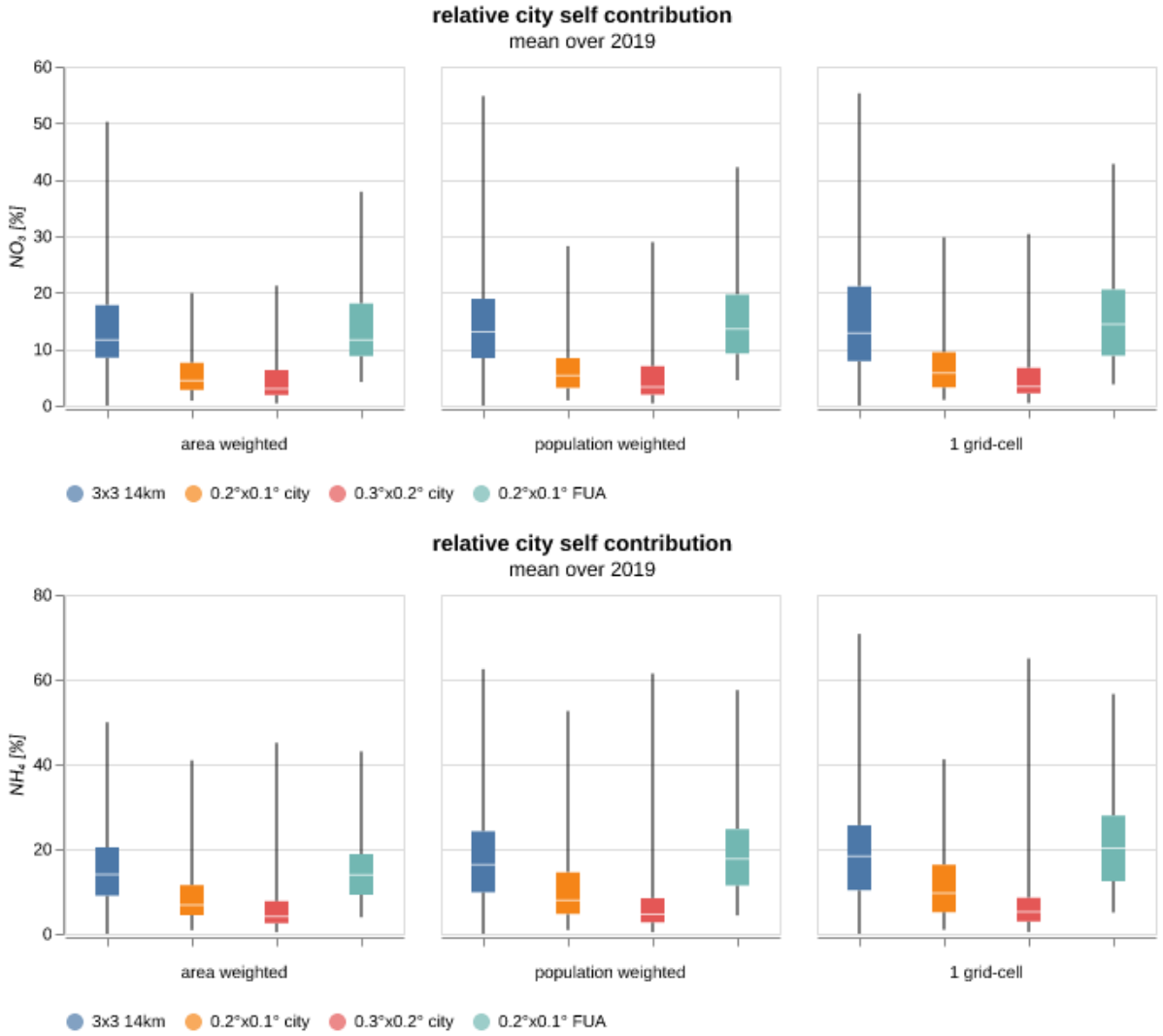


Figure A1 Box quantile diagrams summarizing the local contribution to PM_{10},, primary PM_{10}, SO_4, NO_3 and NH_4. concentrations for the 79 European cities, For each of the 3 plots, the receptor is the same: from left: 1) area weighted receptor = source, 2) population weighted receptor = source and 3) centre grid cell. Within each plot, sources (and model resolution) vary: '3x3 14 km' means that the city source area is defined by 3x3 grid cells in 0.25x0.125 degree model resolution, '0.2x0.1 city' means that core city area is the source and the model is run in 0.2x0.1 degree resolution, '0.3x0.2 city' means that core city area is the source and the model is run in 0.3x0.2 degree resolution and '0.2x0.1 FUA' means that FUA is the source area and the model is run in 0.2x0.1 degree resolution. The box borders show the 10th and 90th percentile and the two extremities are the max and min values. The horizontal line within the box represents the median.

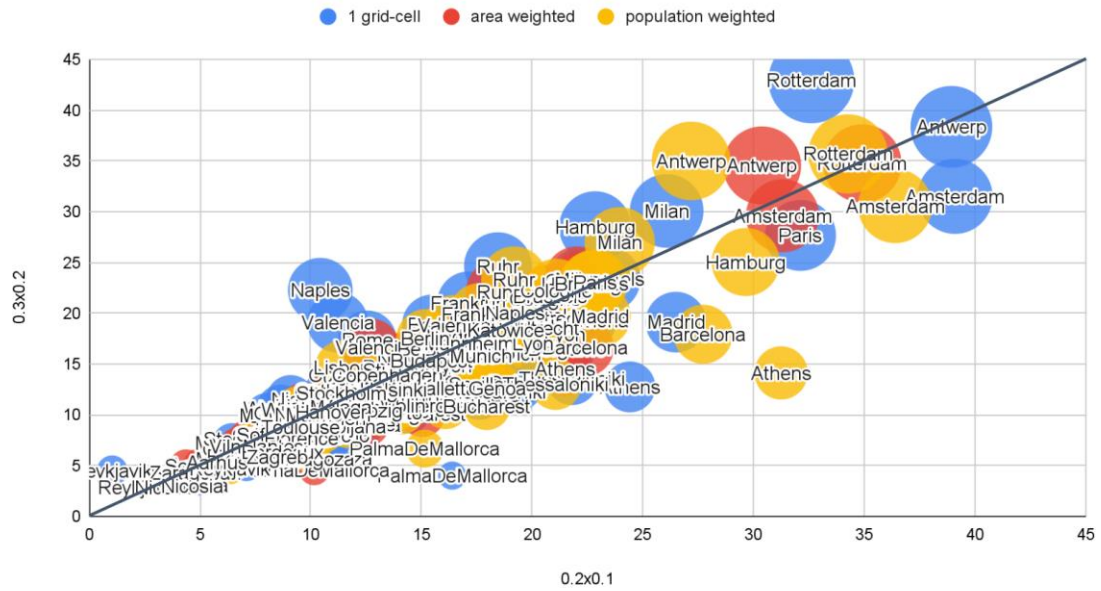


Fig A2: Total contribution/concentration (ug/m3) to NO_2 concentrations in European cities (x-axis for 0.2x0.1 degree resolution, y-axis for 0.3x0.2 degree resolution). The city source areas are the core city area, but the receptor area is either the mid-grid cell, area weighted concentrations for the core city or population weighted concentrations for the core city. The size of the circles represent the total concentrations.

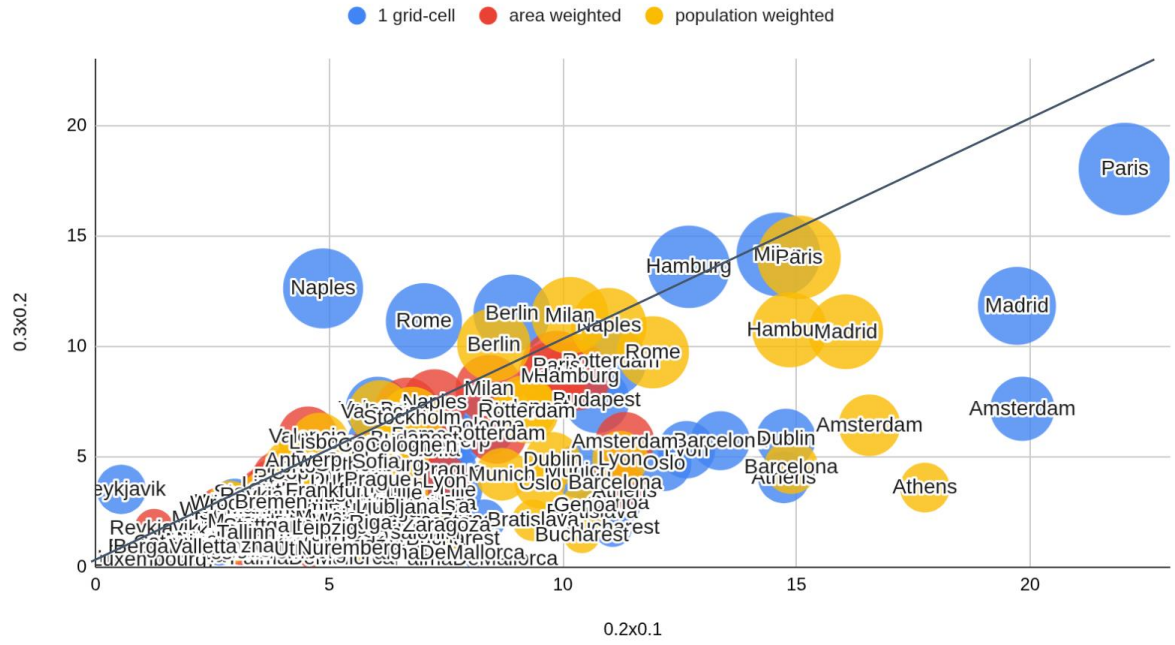


Figure A3: Local contribution (ug/m3) to NO_2 concentrations in European cities (x-axis for 0.2x0.1 degree resolution, y-axis for 0.3x0.2 degree resolution). The city source areas are the core city area, but the receptor area is either the mid-grid cell, area weighted concentrations for the core city or population weighted concentrations for the core city. The size of the circles represent the local contribution (higher contribution means larger circles).

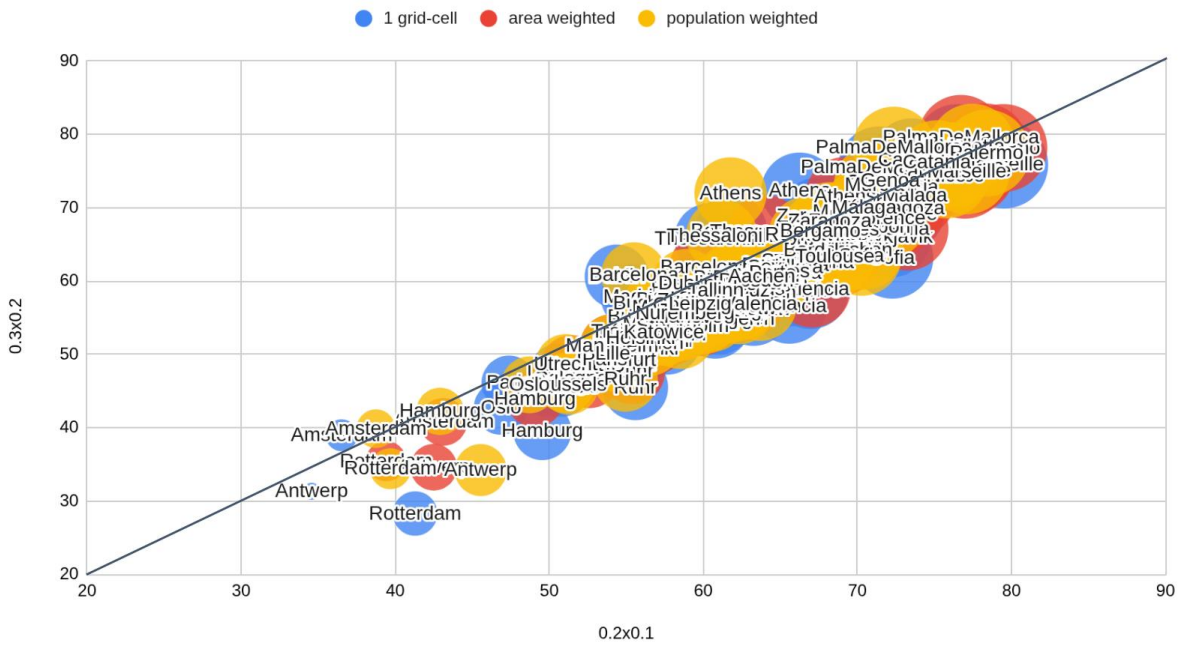


Figure A4: Total contribution/concentration (ug/m3) to O_3 concentrations in European cities (x-axis for 0.2x0.1 degree resolution, y-axis for 0.3x0.2 degree resolution). The city source areas are the core city area, but the receptor area is either the mid-grid cell, area weighted concentrations for the core city or population weighted concentrations for the core city. The size of the circles represent the size of the total concentration.

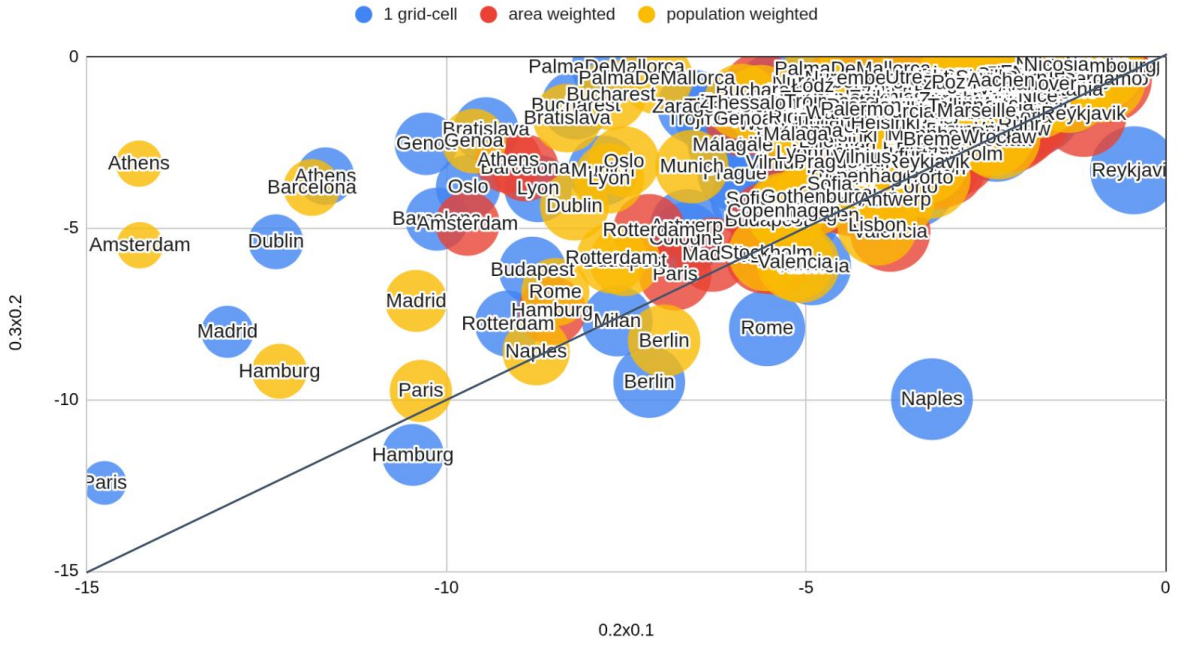


Figure A5: Local contribution (ug/m3) to O_3 concentrations in European cities (x-axis for 0.2x0.1 degree resolution, y-axis for 0.3x0.2 degree resolution). The city source areas are the core city area, but the receptor area is either the mid-grid cell, area weighted concentrations for the core city or population weighted concentrations for the core city. The size of the circles represent the local contribution (larger circles means smaller negative contribution).

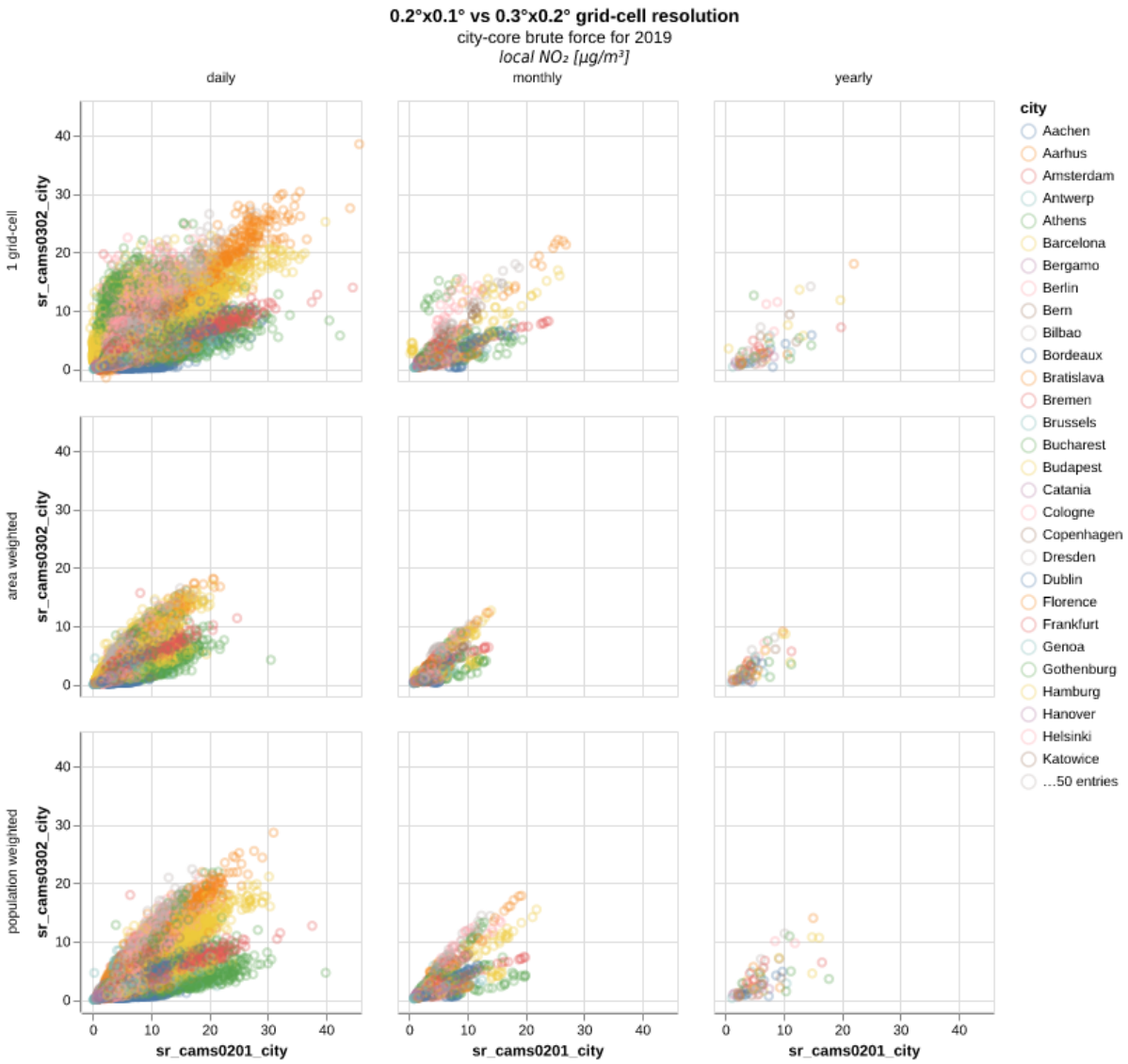


Figure A6: Effect of resolution on local NO_2 contributions (ug/m3). The source area is always core city area, the receptor area is either mid-grid cell, area weighted concentrations in core city or population weighted areas in core city. Left column: Daily data for 2019, middle column: monthly data for 2019, Right column: Yearly average data.

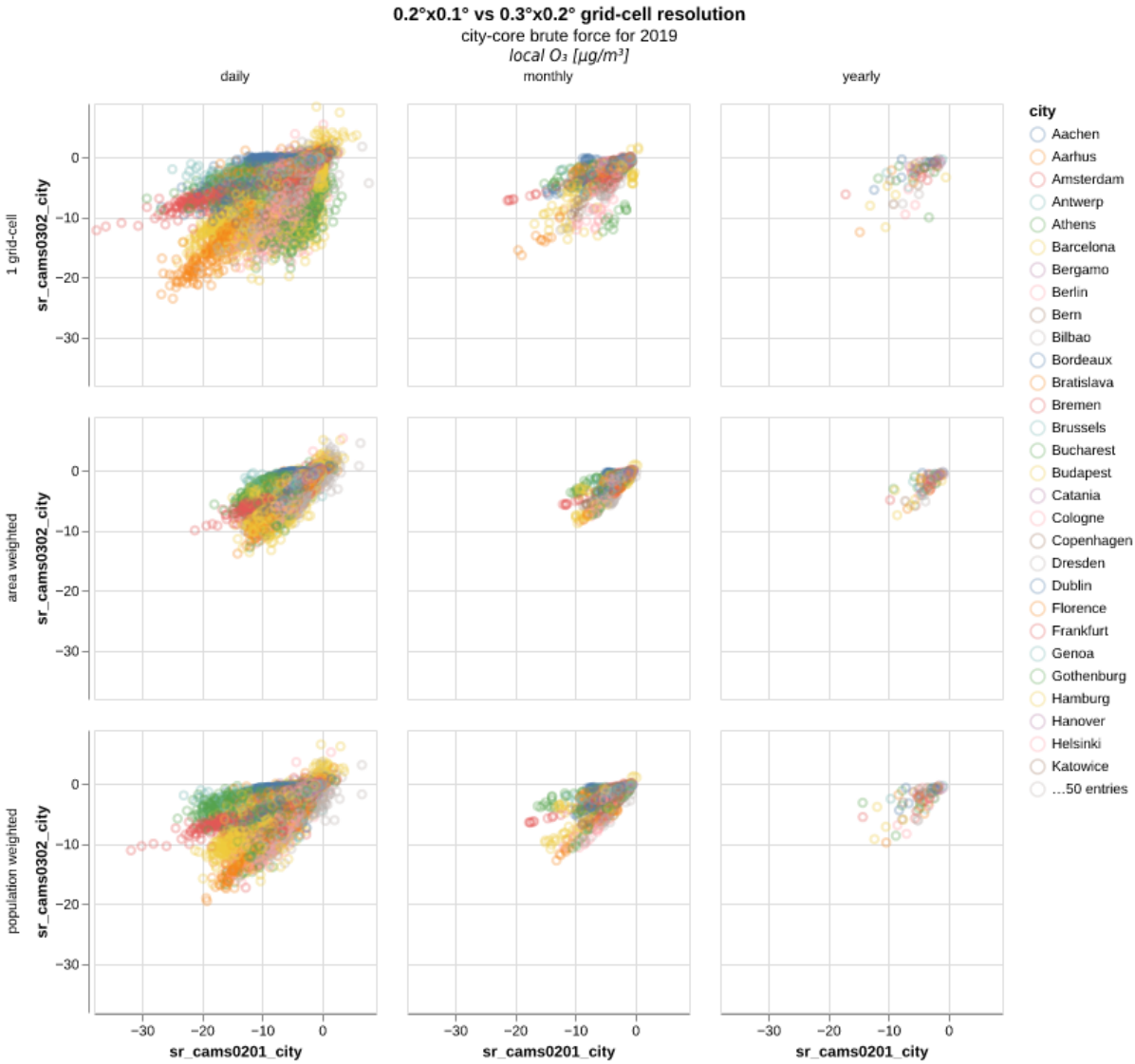


Figure A7: Effect of resolution on local O_3 contributions (ug/m3). The source area is always the core city area, the receptor area is either mid-grid cell, area weighted concentrations in core city or population weighted areas in core city. Left column: Daily data for 2019, middle column: monthly data for 2019, Right column: Yearly average data.

11. APPENDIX B

Table B1: European countries sorted by highest to lowest relative standard deviation for $PM_{2.5}$.

Country Code	Value
MDA	0.58452
TUR	0.524138
MKD	0.505146
UKR	0.469639
MLT	0.454142
ALB	0.452168
SRB	0.447338
KOS	0.44693
RUS	0.410813
MNE	0.409672
EST	0.39876
BIH	0.396314
CYP	0.375146
BGR	0.370363
GRC	0.357632
BLR	0.350079
LVA	0.349286
HRV	0.345444
IRL	0.328036
ROU	0.324135
ISL	0.323057
DNK	0.321274
LUX	0.320765
SVN	0.316865
AUT	0.315645
HUN	0.312726
LTU	0.312338
FIN	0.310828
ITA	0.309557
CZE	0.300207

SVK	0.297181
NLD	0.296421
NOR	0.295081
POL	0.292615
FRA	0.29153
SWE	0.2904
DEU	0.284933
PRT	0.267561
BEL	0.262321
ESP	0.262187
CHE	0.24606
GBR	0.235246

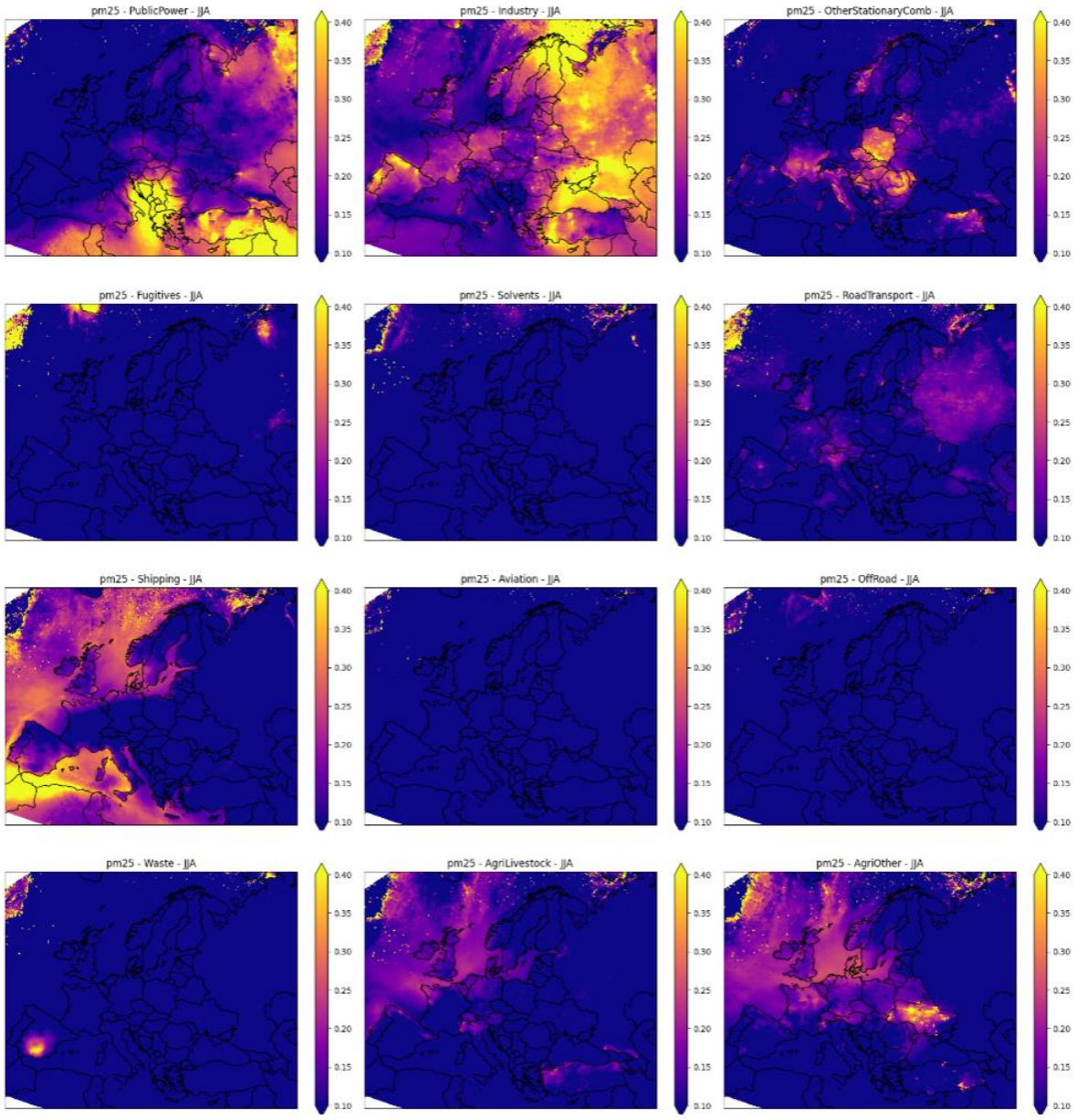


Figure B1: The fractional contribution from each sector to $PM_{2.5}$ in Summer 2019.

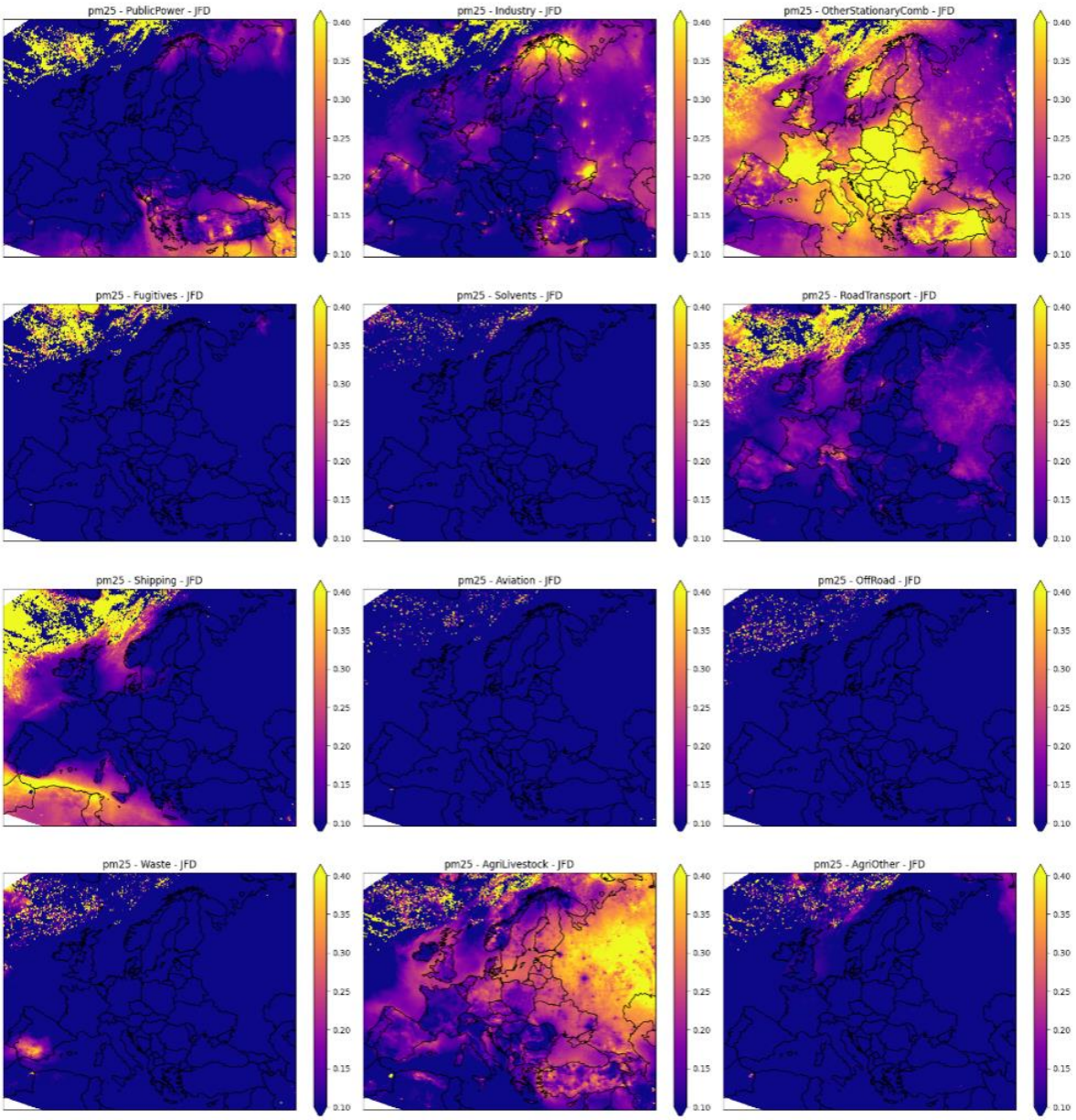


Figure B2: The fractional contribution from each sector to $PM_{2.5}$ in Winter 2019.

12. APPENDIX C

In sections 6.2 and 6.3, we showed that the variability in $PM_{2.5}$ concentrations due to emission uncertainties is often quite significant, reaching up to nearly 70% of the reference concentration — see, for example, Table 6.1 for Bratislava. Furthermore, we observed that this variability tends to be higher when the contribution from the residential sector is predominant. In such cases, the variability is largely driven by uncertainties in the temporal factors rather than by uncertainties in total annual emissions (Figure 6.3).

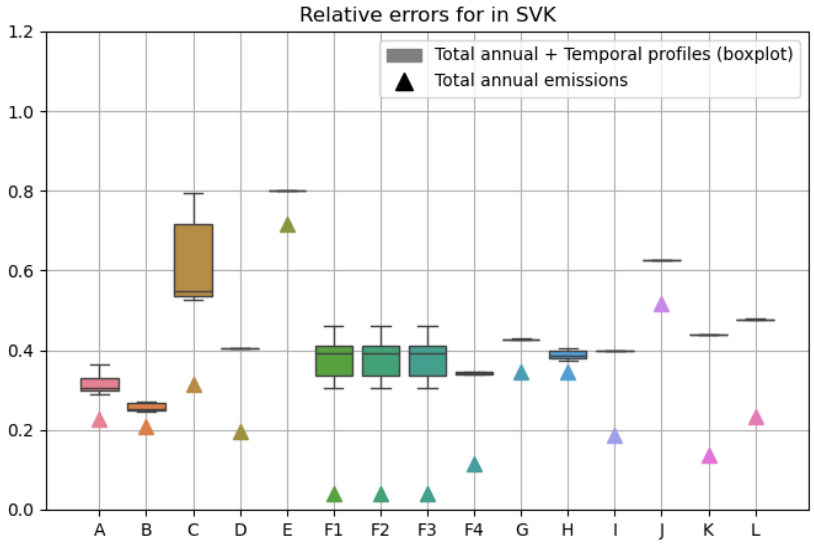
To explore this aspect further, we examined the uncertainties calculated for each GNFR sector prior to their aggregation into ACT sectors. The uncertainties displayed here refer to the pollutants used as proxies to characterize the uncertainty of each sector (see Section 6.2.2 for more details). Figure C1 shows relative hourly uncertainties by GNFR sector for the entire year 2019, for the four countries corresponding to the cities used in the case studies of section 6.3. The triangles show uncertainties derived from annual emissions uncertainties only (Super et al., 2023; 2024), while boxplots represent the full uncertainties i.e. including both uncertainties from annual emissions and the ones associated with the temporal factors (Guevara et al., 2024).

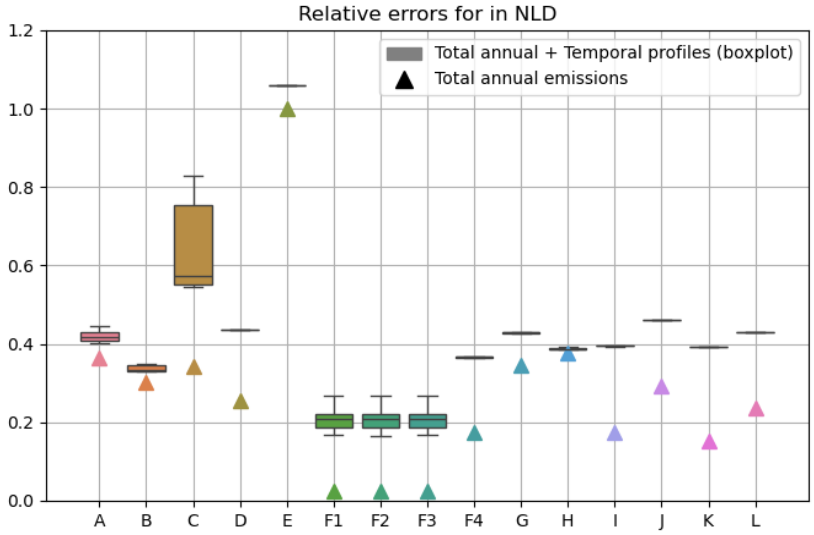
What clearly emerges is that GNFR sector C (Other Stationary Combustion), which corresponds to the ACT residential sector (RH), is among the sectors with the highest uncertainty in all the countries considered. Most values range between 0.5 and 0.8, corresponding to very high relative uncertainties (50–80%).

It can also be observed that, in general, accounting for uncertainties in temporal profiles (box plots in Figure C1) substantially increases the relative uncertainty compared to the case where only annual emissions uncertainties are considered (triangles in Figure C1). This suggests that the contribution to the total uncertainty is strongly due to temporal uncertainties. Furthermore, observing the high errors associated with the residential sector (box plots in Figure C1), it is evident why the concentration uncertainties are very high in cases where the residential sector is dominant (see the case of Bratislava, Frankfurt and Warsaw, in sections 6.3.3, 6.3.4 and 6.3.6).

For the case of Amsterdam (section 6.3.5), one of the most relevant sectors is *Others* (which includes GNFR sectors E, H, I, J, and M). Referring back to Figure C1, it can be observed that the difference between the only annual emissions uncertainty and the total uncertainty is particularly low for sectors E, H, and J. This explains why, for Amsterdam, the difference between the variability associated with total emission uncertainty and that associated with only annual emissions uncertainty is smaller than in the other cities analyzed.

It is also worth noting that the box plots for many sectors are very narrow. This indicates a low variability in the hourly relative error, as previously discussed in Section 6.2.1 and 6.3.1.


Relative errors for in SVK
Total annual + Temporal profiles (boxplot)
Total annual emissions
1.2
1.0
0.8
0.6
0.4
0.2
0.0
A
B
C
D
E
F1
F2
F3
F4
G
H
I
J
K
L


Relative errors for in NLD
Total annual + Temporal profiles (boxplot)
Total annual emissions
1.2
1.0
0.8
0.6
0.4
0.2
0.0
A
B
C
D
E
F1
F2
F3
F4
G
H
I
J
K
L

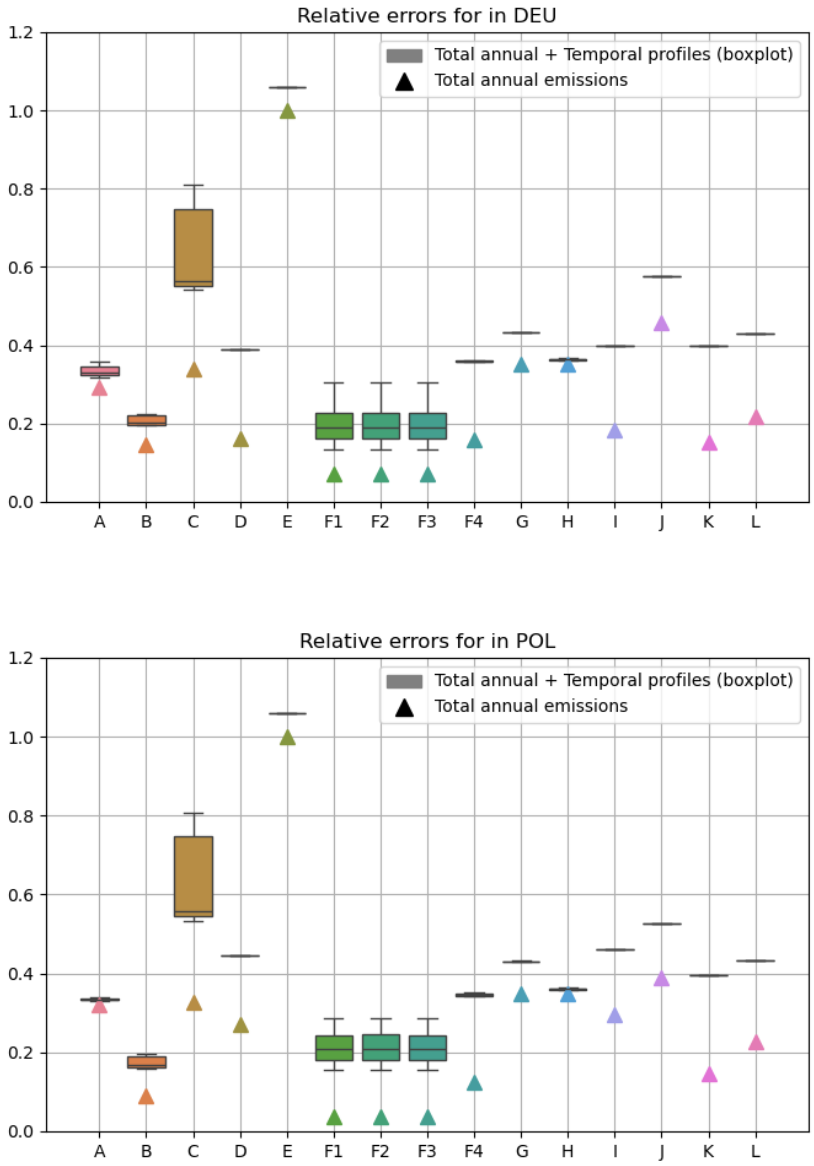
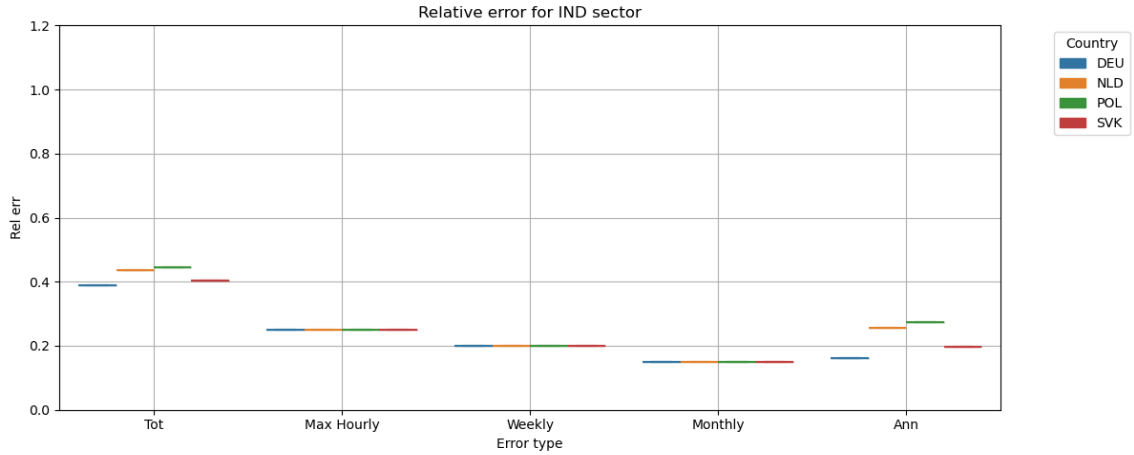
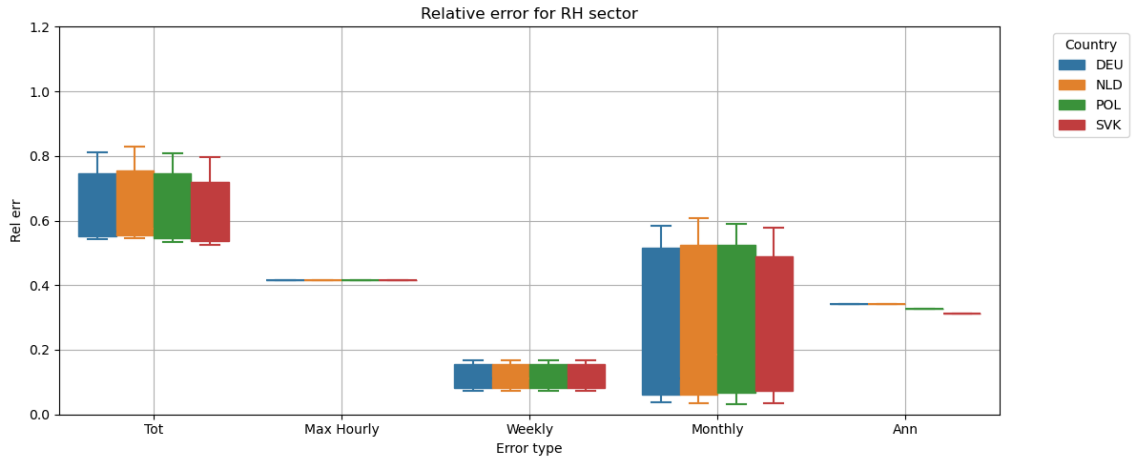


Figure C1 : Hourly relative emission uncertainties by GNFR sector calculated for the entire year 2019. The box plots show uncertainties derived from both temporal factors (Guevara et al., 2024) and annual emission uncertainties (Super et al., 2023; 2024), while triangles represent the only annual emission uncertainties. The uncertainties refer to the four countries corresponding to the cities presented in the case studies of the previous section: Slovakia, the Netherlands, Germany, and Poland (from top to bottom).

As a complement to the analysis, Figure C2 presents the relative emission errors as incorporated in the ACT model (*Tot*) in comparison with their individual components: the relative error associated with the hourly factors, computed as the error corresponding to the daily maximum hourly factor (*Max Hourly*), the relative error from weekly factors (Weekly) and monthly factor (*Monthly*), and the relative error of annual emissions (*Ann*). In other words, *Tot* corresponds to the left-hand side of Equation 6.5, while *Max Hourly*, *Weekly*, *Monthly*, and *Ann* represent the four terms on the right-hand side, after being adapted as inputs for the ACT model. Figure C2 shows these errors for the four countries discussed above (Germany, the

Netherlands, Poland, and Slovakia), aggregated over all hours of the year 2019. It can be observed that for many ACT sectors the relative error associated with hourly factors (*Max Hourly*) is the largest contributor to the total error, while the annual emission uncertainty is the strongest contributor to the Others sector. More generally, the combined contribution of the three temporal factor errors (hourly, weekly, and monthly) exceeds that of the annual emission error in most cases (except for Others and Shipping).





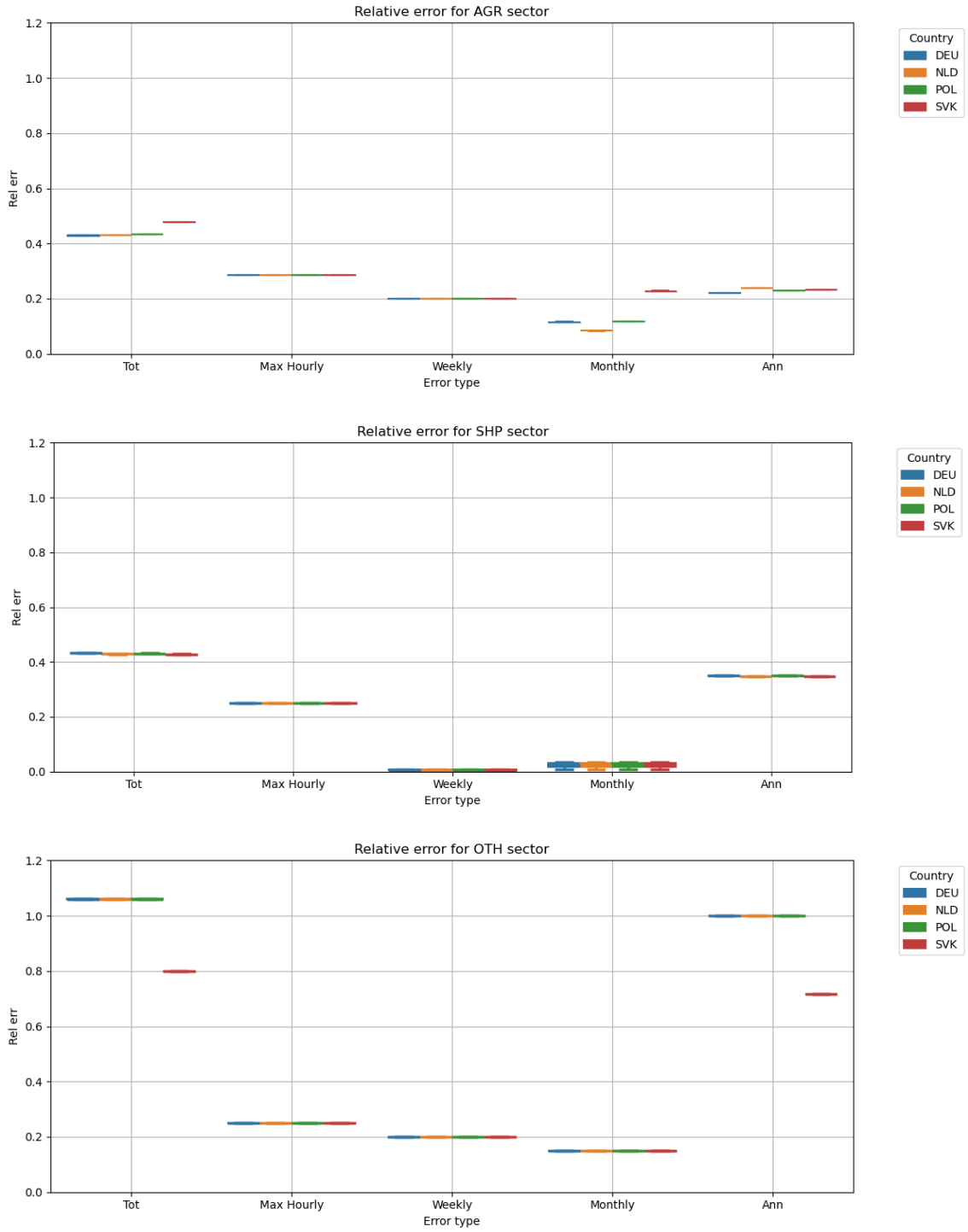


Figure C2 : Relative errors distributed across ACT sectors: *Industry*, *Traffic*, *Residential*, *Agriculture*, *Shipping*, and *Others* (from the top plot to the bottom) for the entire year 2019. Each plot presents the total relative error as injected in ACT (*Tot*), along with its individual components: the relative error associated with the hourly factors, computed as the error corresponding to the daily maximum hourly factor (*Max Hourly*), the relative error associated with the weekly distribution of emissions (*Weekly*), and that related to the monthly distribution (*Monthly*), and finally, the relative error of the annual emissions (*Ann*).

Document History

Version	Author(s)	Date	Changes
1.0	Hilde Fagerli, Lewis Blake, Svetlana Tsyro, Eivind Grøtting Wærsted, Alvaro Valdebenito, Peter Wind, Palmira Messina, Elsa Real, Blandine Raux	11.8.2025	Initial version
1.1	As above	8/9/25	Issued after internal reviews

Internal Review History

Internal Reviewers	Date	Comments
Sam Remy (HYGEOS) Vincent Guidard (MF)	Aug, Sept 2025	Commented on initial version