The Influence of Commuting on Population Exposure to Air Pollution: Toward Global Application With a Proxy on the Degree of Urbanization

Lorenza Gilardi, Thilo Erbertseder, Frank Baier, Heiko Paeth, Tobias Ullmann, and Hannes Taubenböck

Abstract—**Urban populations are significantly affected by air pollution, which poses a major threat to public health. However, standardized and public mobility data, essential for an exposure assessment, are frequently unavailable. Earth observation-derived and model datasets can support large-scale health studies, especially in remote areas with limited data availability. This study investigates the use of a globally derivable variable from remote sensing data to estimate the static versus dynamic population exposure difference. A health risk assessment using a higher and a lower resolution air pollution data was performed. This was achieved by examining air pollution concentrations in two European regions, Lombardy, in Italy, and Germany, incorporating commuting datasets. Accordingly, a retrospective long-term exposure assessment to particulate matter less than 2.5 microns ($PM_{2.5}$), nitrogen dioxide (NO_2), and ozone (O_3) was performed from 2013 to 2022. The study evaluates the difference between the resident and the dynamic population exposed to concentrations exceeding the new limits set by the World Health Organization (WHO). The relation between pollutant concentration and the Fraction of Settlement Area (FSA), a proxy of urbanization levels, derived from the global World Settlement Footprint dataset was explored. Two pollution datasets were used: with European, and global coverage. The analysis decouples daytime and nighttime hours. For each region and pollutant specific FSA thresholds were identified, that maximize the population exposure gap. Our findings highlight the impact of air pollution on population health, revealing widespread exposure exceeding WHO limits, particularly for $PM_{2.5}$, and emphasizing the importance of considering diurnal exposure variations in health risk assessments.**

Index Terms—**Air pollution, remote sensing, risk analysis, urban areas.**

Manuscript received 12 March 2024; revised 5 June 2024; accepted 18 June 2024. Date of publication 1 July 2024; date of current version 24 July 2024. This work was supported by the frame of the German Aerospace Center (DLR)-through the project “Environmental Stressors and Health Costs”. *(Corresponding author: Lorenza Gilardi.)*

Lorenza Gilardi, Thilo Erbertseder, and Frank Baier are with the German Remote Sensing Data Center, German Aerospace Center (DLR), 82234 Weßling, Germany (e-mail: lorenza.gilardi@dlr.de).

Heiko Paeth and Tobias Ullmann are with the Institute for Geography and Geology, University of Würzburg, 97070 Würzburg, Germany.

Hannes Taubenböck is with the German Remote Sensing Data Center, German Aerospace Center (DLR), 82234 Wessling, Germany, and also with the Institute for Geography and Geology, University of Würzburg, 97074 Würzburg, Germany.

This article has supplementary downloadable material available at https://doi.org/10.1109/JSTARS.2024.3420155, provided by the authors.

Digital Object Identifier 10.1109/JSTARS.2024.3420155

I. Introduction

The most recent number of premature deaths in the European Union attributable to air pollution exposure is approximately three hundred thousand per year [1]. According to the estimation given by the World Health Organization (WHO), this number increases to about 4.2 Million on a global level [2]. The exposure to air pollution particularly affects the urban population, whose share in 2022 reached 57% of the global population and 75% in Europe [3]. Moreover, beyond the urban-rural difference in exposure, the air pollution burden is also not equally distributed worldwide. The WHO estimates that 89% of the premature deaths occur in countries classified as low- to middle income, where increasing pollution levels can be observed [2]. Conversely, countries classified as high or upper middle income, although at higher pollution levels, show decreasing air pollution trends over the last decade [4], [5].

Nevertheless, the exposure to an even small concentration of air pollutants poses health risks [6], [7]. In the new Air Quality Guidelines (WHO-AQG) released in 2021, the WHO further decreased the threshold of the maximum air pollutants concentrations considered acceptable in the short- and long-term exposure ranges [8]. For this reason, an accurate estimation of the population exposure to air pollution becomes even more crucial. This information is essential for improving epidemiological studies, deepening our understanding of exposome-response mechanisms and, most importantly, for the accurate planning of active and passive countermeasures to minimize detrimental effects on the population.

In the body of literature on environment and health, exposure indicators are often provided in the form of air pollution and air quality maps. The population, if considered at all, is usually statically linked to the place of residence [9], [10]. Several studies have demonstrated that neglecting the mobility patterns of the population and the pollutants' diurnal variability, leads to errors in the population exposure estimation [11], [12]. In addition, the selected pollution data source significantly affects the exposure assessment due to its specific spatial and temporal coverage and its defined resolutions. For studies conducted at a large scale, a tradeoff in one of these aspects is, due to the current data situation, unavoidable. Especially in remote areas and in the global south, this phenomenon is even more prominent. Datasets with global coverage, such as satellite remote sensing data due to their orbit, have limitations in temporal coverage, while modeled data exhibit shortcomings in spatial resolution. Consequently, the need to perform a health risk and exposure assessment on a large, if not global, scale is intrinsically incompatible with the possibility of having highly detailed information at the individual level of exposed people. The phenomenon is described in epidemiology as "ecological fallacy," which refers to the loss of information in observational studies when inferences about individuals are made based on aggregated group data [13].

The main scope of this work is to propose a methodology to estimate the gap in the measured population exposure to air pollution when the population diurnal mobility is not considered. Target pollutants are: particulate matter smaller than 2.5 micron ($PM_{2.5}$), ozone (O_3), and nitrogen dioxide (NO_2). Particularly, it is investigated whether settlement density can be used to infer the information loss when data on mobility patterns are unavailable. If settlement density turns out to be a satisfactory proxy, comparisons among areas lacking consistent mobility data would become possible.

The settlement density is derived from the globally available World Settlement Footprint dataset. A retrospective long-term exposure assessment to air pollution is performed for the period from 2013 to 2022. We also tested the applicability of globally-available air pollution dataset for the population exposure assessment, in comparison to a regional, higher resolution dataset. These are remote-sensing derived parameters on settlements and air quality that are available worldwide. The study is carried out in two European regions: the Lombardy region in northern Italy and the entire country of Germany. The new concentration exposure limits suggested by the WHO-AQG were consistently used as references [8].

Furthermore, we examined global assessments. The overarching goal of this paper is to evaluate the feasibility of a health risk assessment due to air pollution in remote areas where no mobility patterns are provided and where only coarse-resolution data are available.

The rest of this article is structured as follows. Section II describes the study areas considered, the data sets used, and the methodology adopted. Section III follows outlining the results. Section IV provides a discussion including strengths and limitations of the approach with respect to the current literature as well as the perspectives for future work. Section V concludes this article.

II. Data and Methodology

A. Study Areas

As study areas, we selected the Lombardy region in northern Italy and the entire country of Germany. The choice is based on the following criteria: availability of open access datasets of mobility and air pollution, heterogeneity of degree of urbanization, and air pollution profiles.

Lombardy's land surface covers 23 000 square kilometers. It hosts one fifth of the Italian population, i.e., roughly ten million residents, and its population density in 2022 was of 432.5 inhabitants per square kilometer [14]. This is more than four times the average European population density. Most of the population is concentrated along the Po Valley, which is highly urbanized and industrialized, making it one of the most polluted regions in Europe [15]. Here, poor air quality is often observed due to the peculiar orographic profile that limits the diffusion in the boundary layer [16].

Germany, on the contrary, presents a more scattered urbanization profile, with a population of roughly 84 Million inhabitants spread over about 358 000 square kilometers of land. The country is characterized by major urban centers, where most air pollution sources are located, surrounded by vast low-settlement density areas [17]. Its population density was 235.5 inhabitants per square kilometer in 2022 [14].

B. Data

1) Air Pollution Datasets: Surface concentrations of $PM_{2.5}$, NO_2, and O_3 were derived from the Copernicus Atmosphere Monitoring Service (CAMS) European Air Quality reanalysis dataset. This is a model-based dataset constituted by an ensemble of seven regional chemical transport models (nine, since an update occurred in 2019). Present-day, forecasts, and previous-day analyses of gridded values of pollutants concentrations are delivered daily, with an hourly temporal resolution and a spatial resolution of $0.1^\circ \times 0.1^\circ$ [18], [19]. This results in approximately 840 grid cells covering Lombardy and 7968 covering Germany. Such datasets are available for the continental Europe. A data validation is carried out each year and reanalyses data of the previous year are produced by means of the assimilation of and validation with in-situ observations from the European Environmental Agency.

As an additional source of data, we considered a globally available dataset for this study: the CAMS global reanalysis, Atmospheric Composition Reanalysis 4 (EAC4) data. It is the fourth generation of global reanalysis of atmospheric composition by the European Centre for Medium-Range Weather Forecasts (ECMWF). It combines model data with observations, as chemical transport models are integrated with emissions datasets and atmospheric composition from satellite retrievals [20]. The following retrieval products are included in the model within the timeframe of our study.

1) Global Ozone Monitoring Experiment-2A (GOME-2A, 2007–2015) and GOME-2B (2013–2015)
2) Ozone Monitoring Instrument (OMI) Aura (2004–2015), Microwave Limb Sounder (MLS) Aura (2004–2018)
3) Moderate Resolution Imaging Spectroradiometer (MODIS) Aqua and Terra (2002–2016).

The dataset has a three-hour temporal resolution and a spatial resolution of $0.75^\circ \times 0.75^\circ$, resulting in roughly 25 grid cells for Lombardy and 195 for Germany. Both datasets can be retrieved from the ECMWF Atmosphere Data Store.[1] Since NO_2 and O_3 concentrations are provided as mass ratio units for EAC4, these were converted into mass per volume units by assuming an ambient pressure of 1 atmosphere and a temperature of 25 °C.

[1] [Online]. Available at: https://ads.atmosphere.copernicus.eu.

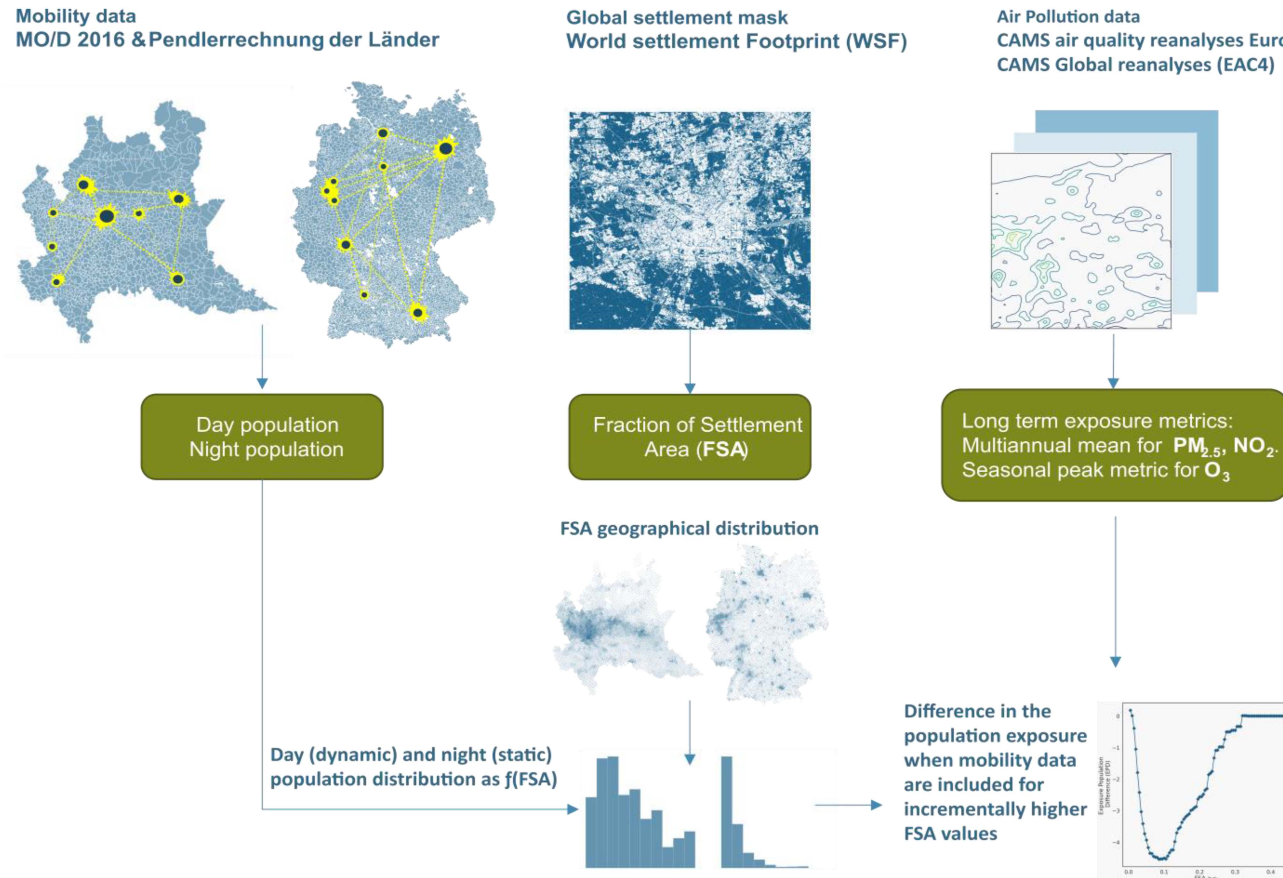

Fig. 1. Graphical representation of the methodology implemented.

2) Mobility Data: Mobility data for Lombardy were obtained from the dataset "Matrice Origine/Destinazione 2016" (MO/D 2016) [21]. This is an improved version of the "Source/Destination Matrix 2014" [22] that was developed within the regional mobility and transport program of Regione Lombardia. The table reports the number of residents, originally derived from the "2011 ISTAT Census", and the hourly commuting within and between each of the 1450 mobility zones into which the region is divided. Data refer to travels due to work, study, occasional, and others, in a typical Monday-Friday working week. The matrix is the result of a complex interaction between transport modeling, online questionnaires, face-to-face interviews, and analysis of available surveys. The spatial subdivision was adopted in the context of a larger study that makes use of mobility data from the MO/D 2016.

Mobility data for Germany were obtained from the "Pendlerrechnung der Länder," a statistic of the Regionaldatenbank Deutschland [23]. The dataset reports the daily mobility between commuting areas in Germany. These correspond to German municipalities or, in the cases of the states of Mecklenburg-Western Pomerania, Rhineland-Palatinate, Schleswig-Holstein, and Thuringia, to cross-municipal associations. The statistic, released yearly, provides the employment-related potential mobility of people who work and/or live in the federal territory of Germany. The result is information on the estimated daily population and the residential one on working days for each of the 6973 commuting areas into which the country is divided.

3) Settlement Data: The World Settlement Footprint 2019 (WSF) was used for this study. It is a global settlement mask with a spatial resolution of ten meters. It is obtained by jointly exploiting multitemporal satellite imagery from Sentinel-1 and Sentinel-2 [24].

C. Methods

A graphical representation of the methodology implemented in this study can be found in Fig. 1.

1) Definition of Timeframe Scenarios: Two timeframes, a day and a nighttime one, were framed based on the hourly population mobility observed for Lombardy in the context of a previous study [11]. There, most travels occurred at the so-called "rush hours" between 6 and 9 A.M. and between 4 and 7 P.M. In the complementary timeframes, the population of the commuting areas remained stable. Therefore, nighttime was defined as the time span between 8 P.M. and 6 A.M. with the population of the commuting areas corresponding to the residential one. 6 A.M. to 8 P.M. is defined as daytime. The same time patterns were applied to Germany.

In this work, we refer to "static population" when considering the residential locations of the population. The so-called

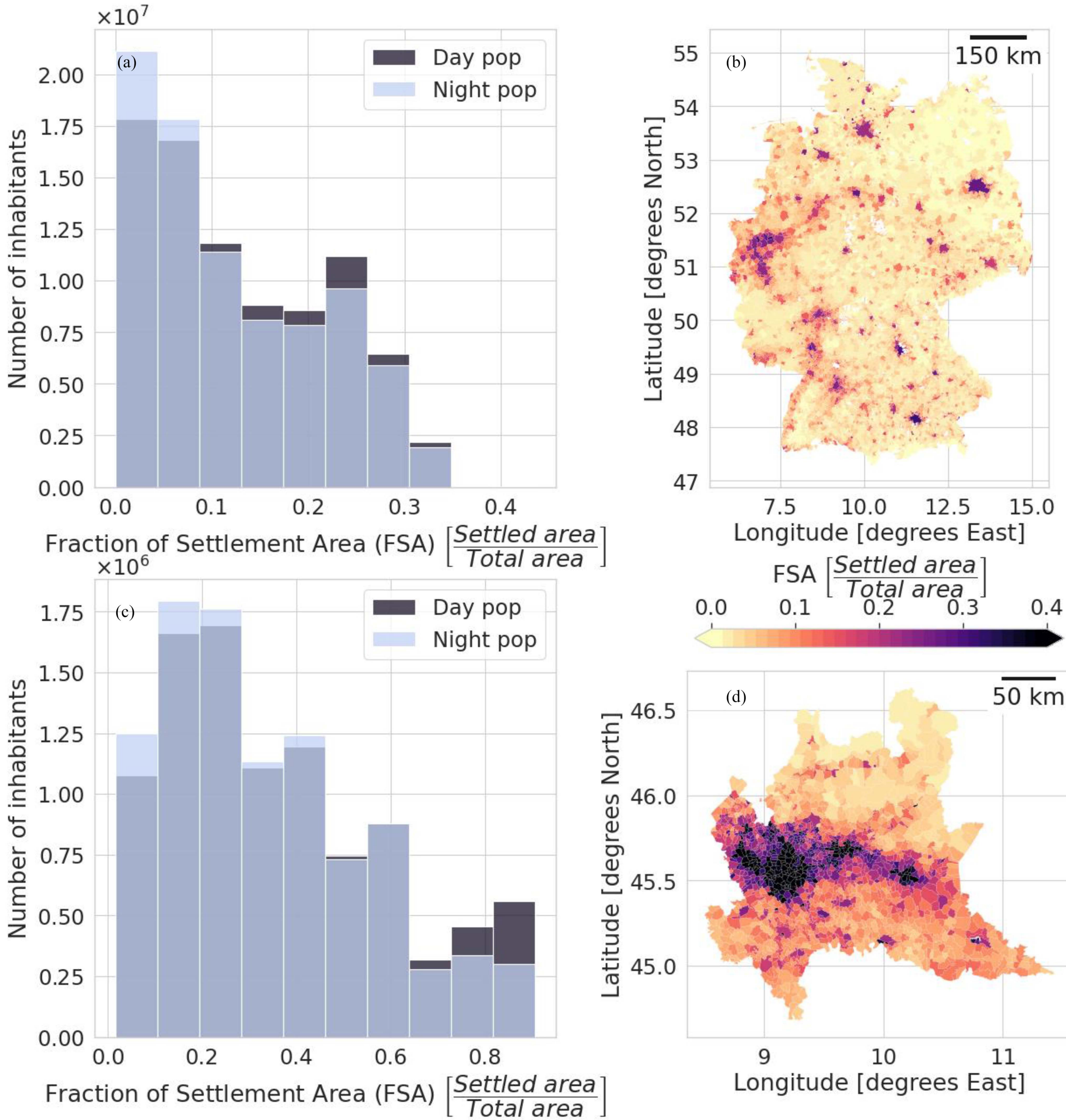

Fig. 2. Static and dynamic population distribution for different Fraction of Settlement Area (FSA) ranges in Germany, sub-figure (a) and Lombardy, sub-figure (c). Geographical distribution of the FSA in the commuting areas of Germany, sub-figure (b), and Lombardy, sub-figure (d).

"dynamic population," referring to the particular whereabouts of the population with respect to the time of the day, was defined as the residential one plus/minus the balance of commuters.

2) Calculation of Pollution Levels at Small Geographical Aggregates: The native resolution of the two datasets utilized, especially the one of EAC4, can be coarser than the size of the small geographical aggregates (i.e., the commuting areas) considered. Therefore, prior to the temporal aggregation, the two datasets were oversampled by increasing the spatial resolution. The respective upscaling factor was chosen by taking the minimum integer allowing the new grid unit to be at least two times smaller than the circle enclosing each commuting area. The interpolation was performed with a bilinear method, as recommended by Stroh et al. [25]. According to their work, the accuracy of coarser datasets resampled bilinearly to a finer grid increases together with the aggregation time. Subsequently, the multiyear mean spatial aggregates were calculated for each vector area for Lombardy and for Germany, separately for day and nighttime hours. O_3 concentrations feature relevant diurnal and seasonal fluctuations. Here, instead of the yearly

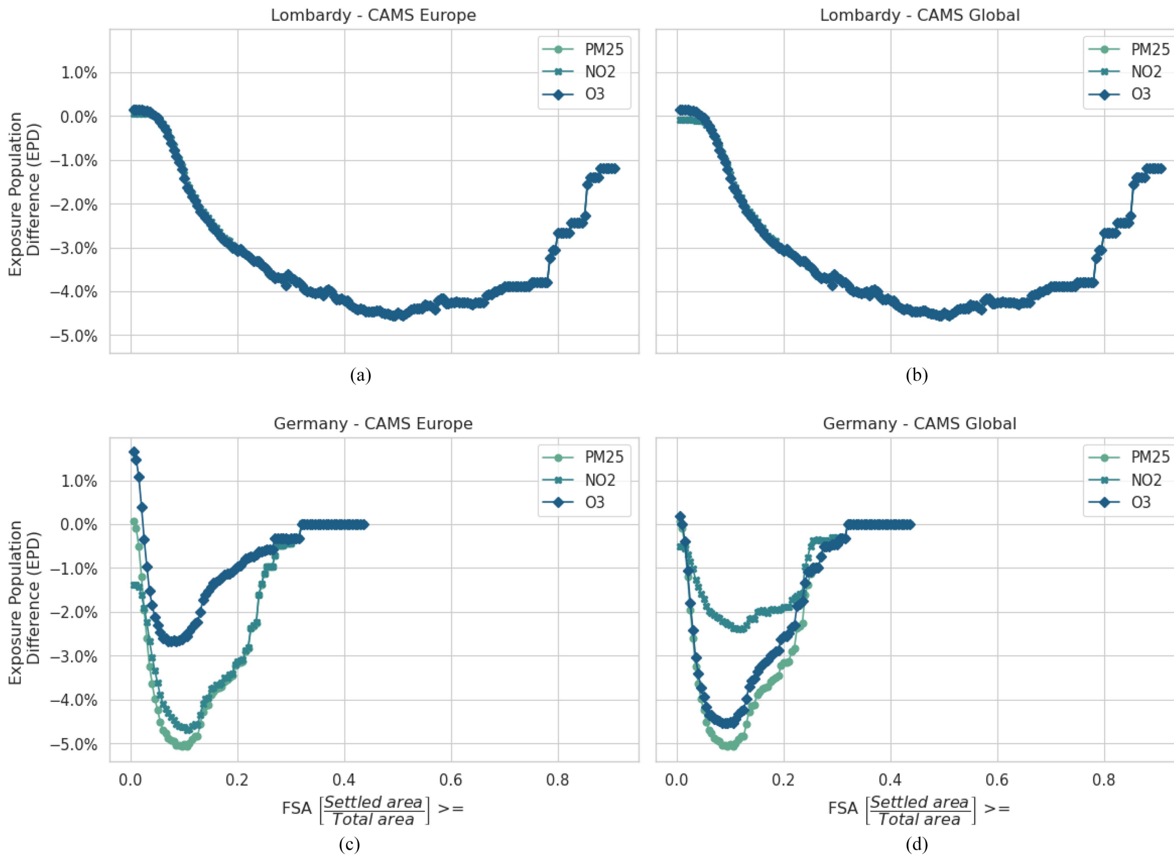

Fig. 3. EPD when considering commuting areas with a FSA surpassing incrementally increasing thresholds. (a) Lombardy using CAMS Europe dataset. (b) Lombardy using the EAC4 dataset. (c) Germany Using the CAMS Europe dataset. (d) Germany using the EAC4 dataset.

mean, the seasonal peak metric was calculated for each year, as recommended by the WHO [8]. This metric was derived by performing the average of daily maxima in an 8-h rolling mean in the six consecutive months of the year presenting the highest six-months running average ozone concentration.

Supplementary Figures 1–6 provide the geographical representation of the long-term concentration metrics for the two regions.

3) Definition of the Fraction of Settlement Area and the Mobility Ratio: It has been shown that the settlement patterns and the degree of urbanization influence the behavior in relation to everyday routines [26], [27]. Against this background, we test in this study, whether the settlement patterns, i.e., the fraction of settlement areas (FSAs) are a feasible proxy for commuting patterns.

The scope is to have a proxy for building density for each of the administrative units considered. The FSA was derived by assigning the value 1 to each pixel marked as “settlement” and the value 0 to those marked as “no settlement” in the WSF layer. Subsequently, the mean over each vector area was derived, resulting in a value comprehended between 0 and 1. The population distribution among different FSA ranges was calculated for Lombardy and for Germany. Two cases were calculated, a static and a dynamic population. This corresponds to the given definition of the timeframes provided previously, i.e., for the assumed daytime and nighttime scenarios (see Section II-B1).

4) Compliance With WHO-AQG: With respect to air pollution exposure, we considered the new limits for long-term exposure as suggested by the WHO-AQG. They correspond to a maximum annual mean of 5 $\mu g/m^3$ for $PM_{2.5}$ and 10 $\mu g/m^3$ for NO_2, and to a maximum of 60 $\mu g/m^3$ of the peak season metric for O_3 [8].

We obtained the shares of the total population exposed to multiannual air pollution concentration aggregates exceeding these values. The population shares were derived considering both, the static and the dynamic population for the daytime and the static population for the nighttime hours. On this basis, we investigated if a gap in the daytime population’s exposure to the three air pollutants could be related to different FSA thresholds. This was expressed in terms of the share of the total population.

For each commuting area, the share of the total population exposed to values exceeding the WHO limits during daily hours were derived. This was based on the static and the dynamic population distributions. The difference of these two values, the exposure population difference (EPD), was evaluated as follows:

$$\text{EPD} = \frac{(\text{Static pop. exposed} - \text{Dynamic pop. exposed})}{\text{Total population}} \quad (1)$$

where the static population corresponds to the residential, i.e., night-time distribution, and the dynamic population corresponds

TABLE I
SHARE OF THE POPULATION EXPOSURE TO POLLUTION CONCENTRATIONS ABOVE THE WHO-AQG, DURING THE DAYTIME AND NIGHTTIME HOURS FOR GERMANY, USING THE CAMS EUROPEAN AND CAMS GLOBAL (EAC4) REANALYSES DATASETS

Germany: CAMS European reanalyses			
Daytime			
	$PM_{2.5}$	NO_2	O_3
Static population (residents)	100%	76.3%	65.9%
Dynamic population (daily population)	99.9%	77.7%	64.2%
Mean conc. of areas Above limits [µg/m3]	10.1	12.4	65.2
Nighttime			
Static population (residents)	100%	86.8%	91.3%
Mean conc. of areas Above limits [µg/m3]	11.1	13.2	68.7

Germany: CAMS global reanalyses (EAC4)			
Daytime			
	$PM_{2.5}$	NO_2	O_3
Static population (residents)	100%	43.6%	85.3%
Dynamic population (daily population)	99.9%	44.1%	85.1%
Mean conc. of areas Above limits [µg/m3]	11.2	10.4	66.2
Nighttime			
Static population (residents)	100%	99.8%	86.3%
Mean conc. of areas Above limits [µg/m3]	13.3	18.9	66.2

to the daily population obtained through the mobility information. “Total population” refers to the overall population of Lombardy and Germany, respectively.

In our analysis, we evaluated the EPD under the condition that the FSA is larger than or equal to a certain threshold and by incrementally increasing the threshold considered. Consequently, we spatially located various areas as hot spots and various amounts of people exposed as a function of the threshold. This process was performed for all three pollutants and for the two different pollution datasets.

III. RESULTS

Fig. 2 shows the distribution of the daytime and the nighttime populations in relation to different FSA ranges. The analysis comprises all commuting areas in Germany and Lombardy that are depicted in the right column presenting the spatial distribution of the FSA.

In both regions, the day population shifts in comparison to the nighttime population towards higher FSA ranges. This means that most of the commuting occurs along the direction of rural-to-urban areas, i.e., areas of higher settlement density.

The main differences between the two regions are the absolute values of the FSA involved. Lombardy features a higher range of FSA values. The FSA reveals areas of very high settlement density with values up to 0.9 whereas the commuting turnover ranges between 0.4 and 0.5 FSA. In contrast, in Germany, the highest FSA is 0.44 and the turnover can be observed around 0.1 FSA. Germany, compared to Lombardy, does not feature a, large, contiguous densely settled area in proportion to the territory extension. It has urban centers scattered across the country, alternating with low-density rural environments, instead. A visible cluster of high settlement density is located in the western part of the country, corresponding to the Rhine-Ruhr metropolitan region, together with a few other major city centers. Lombardy is characterized by an extended triangle-shaped high settlement density area that spans in the central part of the region from west to east, including major cities like Milan, Como, Brescia, and Bergamo. The low-density areas are in the north, where the territory is predominantly mountainous. In the southern part of the territory, there are scattered high-density areas corresponding to bigger cities.

In Tables I and II, we present the share of the total population exposed to long-term air pollutant concentrations exceeding the most recent WHO limits. This is illustrated in the nighttime and the daytime scenarios, where the static and the dynamic populations are considered. Furthermore, the analysis is also performed for the two different air pollution datasets. These calculations consider the totality of commuting areas without distinction of the FSA.

An outstanding finding revealed by almost all combinations is that the majority of the population of both regions is exposed to long-term air pollutant concentrations that exceed the limits suggested by WHO. This is particularly true for $PM_{2.5}$. According to our findings, for all combinations, more than 99% of the population is exposed to excessive concentrations.

For NO_2, a difference in exposure between day- and nighttime is visible, with a higher proportion of the population being exposed at night. This is particularly evident in Germany, where the results remain consistent when one or the other pollution datasets are used. The EAC4 CAMS Global dataset is less capable of detecting NO_2 hotspots compared to the CAMS European reanalyses. When applying EAC4 for Germany, the population exposure share to excessive NO_2 decreases by 32.7%

TABLE II
SHARE OF THE POPULATION EXPOSURE TO POLLUTION CONCENTRATIONS ABOVE THE WHO-AQG, DURING THE DAYTIME AND NIGHTTIME HOURS FOR LOMBARDY, USING THE CAMS EUROPEAN AND CAMS GLOBAL (EAC4) REANALYSES DATASETS

Lombardy: CAMS European reanalyses			
Daytime			
	$PM_{2.5}$	NO_2	O_3
Static population (residents)	99.9%	97.2%	100%
Dynamic population (daily population)	99.7%	97.2%	99.8%
Mean conc. of areas Above limits [µg/m3]	18.8	21	71.3
Nighttime			
Static population (residents)	100%	97.2%	100%
Mean conc. of areas Above limits [µg/m3]	20.1	22.9	79.1

Lombardy: CAMS global reanalyses (EAC4)			
Daytime			
	$PM_{2.5}$	NO_2	O_3
Static population (residents)	100%	95.4%	100%
Dynamic population (daily population)	99.8%	95.5%	99.8%
Mean conc. of areas Above limits [µg/m3]	16.3	12.2	67.5
Nighttime			
Static population (residents)	100%	100%	100%
Mean conc. of areas Above limits [µg/m3]	21	26.2	67.5

(for the static population distribution) and 33.6% (for the dynamic population distribution) during the day. On the contrary, it increases by 13% at night when compared to the results based on the higher resolution CAMS European reanalysis dataset. This also suggests that EAC4 is less sensitive to diurnal variation of NO_2 concentration.

For Germany, the effect of the inclusion of population mobility patterns on the exposure assessment is relatively small. The population exposure, when assuming a static population, is underestimated by 0.5% and 1.4%, when applying EAC4 and CAMS European datasets, respectively.

When using the European dataset in Lombardy, the NO_2 mean concentration is significantly higher than in Germany with an average of 22.9 μg/m^3 over the areas surpassing the WHO-AQG limits. The excessive exposure affects almost the entire population, consistently across daytime, nighttime, and independently from the inclusion of population's mobility patterns. However, when using the EAC4 Global dataset, an increase in the exposures is revealed during nighttime with 26.2 μg/m^3 of mean NO_2 concentration in the areas exceeding the WHO-AQG at night and 12.2 μg/m^3 during the day. This suggests that EAC4 is sensitive to diurnal variation of NO_2 concentration over highly polluted areas.

Concerning O_3, when utilizing the CAMS European dataset, Germany exhibits higher variability in population exposure to the O_3 seasonal peak metric. This variability shifts from approximately 65% exposed people during daytime to about 91% at night. It is important to note that not accounting for population mobility results in a comparatively small difference, leading to an overestimation of 1.7%. When using EAC4, there is little to no diurnal variability or under/overestimation of the exposure share associated with population mobility, with values ranging from 85.1% to 86.3%.

For the Lombardy region, we find that almost the entire population is exposed to excessive concentrations of O_3 in the peak season. This remains consistent for all the calculated scenarios.

In a further step, we evaluated the effects of an incrementally increasing FSA range on the exposed population shares. The results are shown in Fig. 3.

For Germany, the difference in population exposure obtained from static and dynamic population data shows a U-shaped curve for the three pollutants. The turning point in the EPD corresponding to different FSA thresholds varies slightly for the three pollutants. The FSA threshold above which the EPD is minimum, and therefore the underestimation of population exposure in a static scenario is maximum, was 0.11, 0.12, and 0.09 for $PM_{2.5}$, NO_2, and O_3, respectively, using the global dataset. This results in an underestimation of the exposed population corresponding to 5%, 2.4%, and 4.5% of the total one. No exposure overestimation was observed.

For Lombardy, we found the following EPDs: here, we measured a quasi-equal magnitude of EPD for all pollutants, with the minimum measured at an FSA threshold of 0.49. Therefore, when assuming a static population and for commuting areas with an FSA higher than 0.49, 4.5% of Lombardy's population are measured as not exposed to excessive $PM_{2.5}$, NO_2, and O_3 concentrations, while in reality, they are. The average FSA of all geographical aggregates in Lombardy meeting this condition is 0.61. We find that there is little to no overestimation of the population exposure for every FSA threshold. These findings remain consistent when the global and the European datasets are applied.

When using the European data set, the FSA corresponding to the minimum in EPD are 0.11, 0.11, and 0.07 for $PM_{2.5}$, NO_2, and O_3, respectively. The corresponding EPDs are 5%, 4.7%, and 2.6%. An overestimation of exposure in the static scenario, affecting 1.6% of the total population, is observed when considering all commuting areas up to those with FSA values higher or equal to 0.03.

IV. Discussion

In this article, we investigated the influence of commuting on the measured exposure to air pollution. Furthermore, we investigated whether a global health exposure and risk assessment is meaningful, if mobility patterns of the population are inexistent and a proxy information on the degree of urbanization is applied. We performed a retrospective long-term exposure assessment to $PM_{2.5}$, NO_2, and O_3 in relation to the WHO-AQG. This was evaluated for the Lombardy region in Italy and for the entire Germany. The analysis was conducted using the globally available and remote sensing-derived variable FSA. Two sets of air pollution data with different spatial coverage and resolution were used to assess the proportion of populations exposed to concentrations above the WHO-AQG threshold values: The CAMS European reanalysis, available for continental Europe and with a resolution of $0.1° \times 0.1°$, and the EAC4, globally available and with a resolution of $0.75° \times 0.75°$. We applied a static population distribution and a dynamic population distribution, assigned to all commuting areas in the two regions.

In general, we find that Lombardy, as a densely urbanized area, is consistently more affected by high air pollution levels. Regarding the pollution datasets, only small differences were observed between the exposure assessments, as in both cases, the majority of the population was found to be exposed to excessive concentrations of air pollutants. For Germany, however, major differences were observed in the exposure to NO_2 and O_3.

By separating air pollution concentrations into daytime and nighttime periods, the analysis allows us to distinguish between exposures that occur between typical daily commuting peaks and those that occur at night. By accounting for these temporal variations in exposure, the study provides insights into the relationship between population mobility and long-term air pollution exposure, particularly in urban areas where commuting patterns have a significant impact on daily pollution levels.

Our findings on the share of the total population exposed to air pollution in the daytime hours, when using the European dataset, are consistent with the findings of Beloconi and Vounatsou [28]. Here 100% (99.6–10) and 78% (70.2–84.9) of the total German population is measured to be exposed to air pollution levels higher than the $PM_{2.5,}$ and NO_2 limits in 2021. Beloconi et al. provide data for Italy as a whole country, making a direct comparison with single regions like Lombardy difficult. Nevertheless, they reported that 99.7% of the Italian population is exposed to excessive concentrations of $PM_{2.5}$. The introduction of the new stringent limit of 5 μg/m^3, results in a noncompliance across all considered areas, independently from the used dataset. Given that the long-term relative risk of mortality due to all causes provided by the WHO-AQG is 1.08 (1.06–1.09) for an incremental increase of 10 μg/m^3 of $PM_{2.5}$, the exposed population faces an increased risk of mortality, of 4% (3%–4.5%), as a minimum. If a linear dose-response relationship is assumed, as suggested by WHO, this percentage increases incrementally with the $PM_{2.5}$ long-term mean concentration recorded for each commuting area.

For any study area-dataset combination, except for the Lombardy—EAC4 combination, the nighttime population exposure to NO_2 is reported higher. The highest diurnal difference is measured for Germany when using the EAC4 data. This result may be attributable to the accumulation effect of NO_2 at night, especially in the winter months, when photolysis processes are limited to few hours. The lifetime of NO_x at the considered latitude increases in fact from about 3 h in the summer season to about 4.3 h in the winter one [29]. Moreover, it is expected to further increase in cleaner, rural areas where background concentration prevails over emission sources but may still exceed the limit of 10 μg/m^3.

Concerning the O_3 seasonal peak metric, the results obtained using EAC4 show no general differences in day- and nighttime exposures. The results show that it is affecting a large part of the population (around 85% in Germany and almost the entire population in Lombardy). This may be linked to the fact that the global dataset is less capable of capturing the diurnal cycle of O_3 production. Furthermore, the EAC4 validation report, confirms a positive "Modified normalized mean bias" against AirBase background rural observations [30]. This bias between 2003 and 2022 over central Europe is up to 15% for the months from May to December.

The difference between static and dynamic population exposure in the totality of commuting areas appears to be of limited significance when expressed as percentages. However, in absolute numbers, this is translated in a total population exposure gap to $PM_{2.5}$, NO_2, and O_3 for the two regions, respectively, of +0.103, -1.162, and +1.431 million people when using the CAMS European reanalyses dataset and of +0.103, −0.425, and +0.186 million people when using EAC4. Positive values indicate an overestimation of the population exposure and negative values indicate an underestimation when mobility data is not considered. All the above-mentioned considerations are made on the total commuting areas without differentiation by degree of urbanization. The small percentage discrepancy and the exposure underestimation in the dynamic scenario, especially for $PM_{2.5}$, is likely due to the underestimation of population exposure in more urbanized areas being outweighed by the overestimation of exposure in less urbanized areas. Moreover, the new limits suggested by WHO are easily surpassed also in less urbanized and cleaner areas. Therefore, while the total exposed population remains consistent with previous research studies, the lack of differentiation due to location, i.e., highly urbanized versus rural areas, masks the existing local differences in exposure.

The study provides FSA thresholds for the commuting areas at which the exposure gap between a static and a dynamic population scenario is maximum. These thresholds vary, spanning approximately around 0.1 for Germany and around 0.5 for Lombardy. When only commuting areas above this degree of urbanization threshold are considered, 4.6, 4.35, and 2.6 million people are exposed to excessive levels of $PM_{2.5}$, NO_2, and O_3, respectively, when using the CAMS European reanalyses dataset and a static population approach. When using EAC4 pollution data, these numbers become 4.6, 2.44, and 4.18 million. Given that long-term exposure to air pollutants is linked to an increased probability of numerous health outcomes, the social impact of the misclassification of population exposure is considerable. Also, in economic terms, this may lead to unforeseen costs and burden for the public healthcare systems [31].

The study focuses on an approach that provides information at small geographical aggregates, which is particularly useful for studies adopting an ecological approach [32]. This allows for the investigation of health effects in epidemiology for large areas and for an easy integration of additional data. To support ecological studies on the urban scale [33], however, the method would require further refinements. Nevertheless, the inclusion of pollution diurnal variability and mobility patterns aims at mitigating the ecological fallacy effect even if this cannot be completely excluded. The commuting data, in fact, cannot provide information down to the individual level as they represent only the typical mobility patterns in a Monday-to-Friday working week. Data for Germany, in particular, rely exclusively on the information of the location of study/work. They show, therefore, the potential trips, without considering, for example, major changes in mobility behavior that occurred during and after the COVID-19 pandemic [34], [35]. Moreover, the lack of standardized definitions for commuting areas between Lombardy and Germany introduces a potential inconsistency in the analysis across the two regions. For Lombardy, the definition of commuting areas was nonstandard and meant to be optimized for the use of mobility data, whereas in Germany, standard administrative areas are employed. For this reason, big metropolitan areas like Munich and Berlin are considered as a large single geographical unit. The consequence is a loss of granularity in the information after data aggregation.

The subdivision of exposure assessments into day- and nighttime hours is influenced by the habits and activity patterns specific to each country and population. This limits the replicability of our findings to regions with similar sociocultural characteristics. In order to replicate the study in diverse areas, additional local data should be reviewed in order to substantiate the applicability of our approach.

V. Conclusion

A. Summary and Main Findings

The study delivers a numerical assessment of long-term population exposure to air pollutants in two European regions with a total population of more than ninety million individuals. It reveals that the majority of the population is exposed to concentrations of $PM_{2.5}$, NO_2, and O_3 exceeding the limits recommended in the WHO-AQG. Notably, almost the entire population is exposed to excessive $PM_{2.5}$ levels. The study highlights the importance of mobility data in conducting comprehensive health risk assessments, demonstrating substantial differences in population exposure during daytime and nighttime, particularly in Germany where exposure to NO_2 and O_3 is higher at night. In addition, a comparative use of the globally available dataset for the air pollutants (EAC4) was tested, with encouraging results for future applications in remote areas. It was found that, considering the stringent concentration limits proposed by WHO, EAC4 constitutes a viable first option to perform a preliminary exposure and health risk assessment.

B. Major Limitations

One limitation of the study is represented by the lower sensitivity of the globally available air pollution dataset (CAMS reanalysis global, EAC4) to variations between daytime and nighttime concentrations, and differences in concentration magnitudes, when compared to the higher resolution CAMS reanalysis for Europe. Future work could benefit from the use of satellite-based air pollution data to mitigate this limitation [36].

In addition, the study finds no general rule for determining a key FSA threshold linked to the maximum gap in the exposure assessment when mobility data are not available, that could be applicable worldwide. This indicates the need for further research incorporating additional globally available data layers. For example: socioeconomic indicators, and the three-dimensional structure of the built environment [37]. Moreover, additional data would be required to perform a sensitivity analysis in order to assess the possible biases between the mobility datasets of the two analyzed areas.

C. Advantages for the Scientific Community

The study's approach, which includes the use of globally available air pollution data and the FSA, enhances the applicability of its findings beyond the examined regions. This facilitates potential comparisons with other geographical areas worldwide, extending the scope and relevance of this research. The use of the globally available dataset (EAC4) for preliminary health risk assessment can be particularly beneficial in areas where other data sources are not consistently available. The insights provided by the study on the long-term exposure of the population to air pollution can support policymakers and researchers in better addressing air quality management and health challenges in Lombardy and Germany. Furthermore, the study's methodology of decoupling daytime and nighttime exposure offers an increased understanding of population exposure, which could be valuable in developing targeted air quality interventions.

Acknowledgment

The authors acknowledge the free use of products of the Copernicus Atmospheric Monitoring Service.

References

[1] European Environmental Agency, "Europe's air quality status 2023," 2023. Accessed: May 31, 2023. [Online]. Available: https://www.eea.europa.eu/publications/europes-air-quality-status-2023

[2] World Health Organization, "Ambient (outdoor) air pollution," 2024. Accessed: Feb. 02, 2024. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health

[3] The World Bank, "Urban population (% of total population)," 2021. Accessed: Dec. 12, 2022. [Online]. Available: https://data.worldbank.org/indicator/SP.URB.TOTL.IN.ZS?end=2020&locations=EU&start=1960&view=chart

[4] T. Erbertseder et al., "Earth observation-based analysis of NO_2 pollution and settlement growth in megacities," in *Joint Urban Remote Sens. Event (JURSE)*, 2023, pp. 1–4, doi: 10.1109/JURSE57346.2023.10144190.

[5] P. M. Mannucci and M. Franchini, "Health effects of ambient air pollution in developing countries," (in English), *Int. J. Environ. Res. Public Health*, vol. 14, no. 9, Sep. 2017, Art. no. 1048, doi: 10.3390/ijerph14091048.

[6] M. Strak et al., "Long term exposure to low level air pollution and mortality in eight European cohorts within the ELAPSE project: Pooled analysis," (in English), *Brit. Med. J.*, vol. 374, Sep. 2021, Art. no. 1904, doi: 10.1136/bmj.n1904.

[7] M. Stafoggia et al., "Long-term exposure to low ambient air pollution concentrations and mortality among 28 million people: Results from seven large European cohorts within the ELAPSE project," (in English), *Lancet Planet. Health*, vol. 6, no. 1, pp. E9–E18, Jan. 2022. [Online]. Available: http://WOS:000744596400006

[8] World Health Organization, "WHO global air quality guidelines. Particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide," Geneva: World Health Organization, 2021. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK574594/

[9] R. Beelen et al., "Effects of long-term exposure to air pollution on natural-cause mortality: An analysis of 22 European cohorts within the multicentre ESCAPE project," (in English), *Lancet*, vol. 383, no. 9919, pp. 785–795, Mar. 2014, doi: 10.1016/S0140-6736(13)62158-3.

[10] N. Künzli et al., "Ambient air pollution and atherosclerosis in Los Angeles," (in English), *Environ. Health Perspectives*, vol. 113, no. 2, pp. 201–206, Feb. 2005, doi: 10.1289/ehp.7523.

[11] L. Gilardi, M. Marconcini, A. Metz-Marconcini, T. Esch, and T. Erbertseder, "Long-term exposure and health risk assessment from air pollution: Impact of regional scale mobility," (in English), *Int. J. Health Geographics*, vol. 22, no. 1, May 19, 2023, Art. no. 11, doi: 10.1186/s12942-023-00333-8.

[12] Y. M. Park and M. P. Kwan, "Individual exposure estimates may be erroneous when spatiotemporal variability of air pollution and human mobility are ignored," (in English), *Health Place*, vol. 43, pp. 85–94, Jan. 2017, doi: 10.1016/j.healthplace.2016.10.002.

[13] John J. Hsieh, "Ecological fallacy," in *Encyclopedia Britannica*, Sep. 2017. Accessed: Feb. 4, 2024. [Online]. Available: https://www.britannica.com/science/ecological-fallacy

[14] Eurostat., "Population density by NUTS 3 region," European Commission, Eurostat, 2024. [Online]. Available: http://data.europa.eu/88u/dataset/gngfvpqmfu5n6akvxqkpw

[15] Eurostat., "Regions in Europe - 2021 interactive edition," 2022. Accessed: Sep. 20, 2022. [Online]. Available:https://ec.europa.eu/eurostat/cache/digpub/regions/#total-population

[16] A. Maurizi, F. Russo, and F. Tampieri, "Local vs. external contribution to the budget of pollutants in the Po Valley (Italy) hot spot," *Sci. Total Environ.*, vol. 458–460, pp. 459–465, Aug. 1, 2013, doi: 10.1016/j.scitotenv.2013.04.026.

[17] H. Taubenböck et al., "To be, or not to be 'urban'? A multi-modal method for the differentiated measurement of the degree of urbanization," *Comput., Environ. Urban Syst.*, vol. 95, Jul. 1, 2022, Art. no. 101830, doi: https://doi.org/10.1016/j.compenvurbsys.2022.101830.

[18] F. Meteo et al., "CAMS European air quality forecasts, ENSEMBLE data," 2024. Accessed: Oct. 2022, 2021. [Online]. Available: https://confluence.ecmwf.int/display/CKB/CAMS+Regional%3A+European+air+quality+reanalyses+data+documentation

[19] D. Akritidis et al., "A complex aerosol transport event over Europe during the 2017 Storm Ophelia in CAMS forecast systems: Analysis and evaluation," *Atmos. Chem. Phys.*, vol. 20, no. 21, pp. 13557–13578, 2020, doi: 10.5194/acp-20-13557-2020.

[20] ECMWF, "Global Reanalysis," 2024. Accessed: Jun. 06, 2023. [Online]. Available: https://confluence.ecmwf.int/display/CKB/CAMS%3A+Reanalysis+data+documentation

[21] Regione Lombardia Open Data, "Matrice OD2016 - Passeggeri," 2023. [Online]. Available: https://www.dati.lombardia.it/Mobilit-e-trasporti/Matrice-OD2016-Passeggeri/tezw-ewgk/about_data

[22] Regione Lombardia, "Nota metodologica utilizzo dati MatriceO/D," 2014. [Online]. Available: https://www.regione.lombardia.it/wps/wcm/connect/d598a97a-05bf-4de6-8822-4213f20e8271/Nota+metodologica+Matrice+OD.pdf?MOD=AJPERES&CACHEID=ROOTWORKSPACE-d598a97a-05bf-4de6-8822-4213f20e8271-lCXZLd0

[23] L. Pendlerrechnung der, "193 Länderstatistiken im Bereich Arbeitsmarkt," 2023. [Online]. Available: https://www.regionalstatistik.de/genesis/online/

[24] D. Palacios-Lopez et al., "New perspectives for mapping global population distribution using world settlement footprint products," (in English), *Sustainability-Basel*, vol. 11, no. 21, Nov. 2019, Art. no. 6056, doi: 10.3390/su11216056.

[25] E. Stroh, L. Harrie, and S. Gustafsson, "A study of spatial resolution in pollution exposure modelling," *Int. J. Health Geograph.*, vol. 6, Jun. 4, 2007, Art. no. 19, doi: 10.1186/1476-072X-6-19.

[26] F. Wen, Y. Jiang, and L. Jiang, "Intercity mobility pattern and settlement intention: Evidence from China," *Comput. Urban Sci.*, vol. 2, no. 1, Dec. 14, 2022, Art. no. 46, doi: 10.1007/s43762-022-00075-6.

[27] V. M. Carlow, O. Mumm, D. Neumann, N. Schmidt, and T. Siefer, "TOPOI Mobility: Accessibility and settlement types in the urban rural gradient of Lower Saxony – opportunities for sustainable mobility," *Urban, Plan. Transport Res.*, vol. 9, no. 1, pp. 207–232, Jan. 1, 2021, doi: 10.1080/21650020.2021.1901603.

[28] A. Beloconi and P. Vounatsou, "Revised EU and WHO air quality thresholds: Where does Europe stand?," (in English), *Atmos. Environ.*, vol. 314, Dec. 1, 2023, Art. no. 120110, doi: 10.1016/j.atmosenv.2023.120110.

[29] K. Lange, A. Richter, and J. P. Burrows, "Variability of nitrogen oxide emission fluxes and lifetimes estimated from Sentinel-5P TROPOMI observations," (in English), *Atmos. Chem. Phys.*, vol. 22, no. 4, pp. 2745–2767, Mar. 1, 2022, doi: 10.5194/acp-22-2745-2022.

[30] European Environment Agency, "[DEPRECATED] AirBase - the European air quality database," 2022. [Online]. Available: http://data.europa.eu/88u/dataset/data_airbase-the-European-air-quality-database-8

[31] OECD, *The Economic Consequences of Outdoor Air Pollution*, 2016.

[32] D. Weismann, M. Möckel, H. Paeth, and A. Slagman, "Modelling variations of emergency attendances using data on community mobility, climate and air pollution," (in English), *Sci. Rep.*, vol. 13, no. 1, Nov. 23, 2023, Art. no. 20595, doi: 10.1038/s41598-023-47857-4.

[33] S. K. T. Iungman et al., "The impact of urban configuration types on urban heat islands, air pollution, CO_2 emissions and mortality in Europe: An ecological analysis," *Lancet Planet. Health*, pp. E489–E505, 2024.

[34] L. Mejía-Dorantes, L. Montero, and J. Barceló, "Mobility trends before and after the pandemic outbreak: Analyzing the metropolitan area of Barcelona through the lens of equality and sustainability," (in English), *Sustainability-Basel*, vol. 13, no. 14, Jul. 2021, Art. no. 7908, doi: 10.3390/su13147908.

[35] C. Balbontin, D. A. Hensher, and M. J. Beck, "Relationship between commuting and non-commuting travel activity under the growing incidence of working from home and people's attitudes towards COVID-19," (in English), *Transportation*, Jul. 1, 2023, doi: 10.1007/s11116-023-10403-2.

[36] J. Handschuh, T. Erbertseder, and F. Baier, "On the added value of satellite AOD for the investigation of ground-level PM2.5 variability," *Atmos. Environ.*, vol. 331, Aug. 18, 2024, Art. no. 120601, doi: https://doi.org/10.1016/j.atmosenv.2024.120601.

[37] T. Esch et al., "World Settlement Footprint 3D-A first three-dimensional survey of the global building stock," (in English), *Remote Sens. Environ.*, vol. 270, Mar. 1, 2022, Art. no. 112877, doi: 10.1016/j.rse.2021.112877.

Lorenza Gilardi received the Master of Science degree in environmental and land planning engineering from Politecnico di Milano, Milan, Italy, in 2016.

In 2015, she worked as a Research Engineer with the Department of Environmental Chemistry, the Royal Institute of Technology, Stockholm, Sweden. From 2018 to 2019, she worked as a Scientific Employee with the Department of Analytical Chemistry, the Technical University of Munich, Munich, Germany. Since 2019, she has been employed as a Researcher and Project Manager with the German Remote Sensing Data Center, the German Aerospace Center, Department "Atmosphere". From 2023, she is being enrolled as Ph.D. student at the Faculty of Geography, the Julius-Maximilian University of Würzburg, Würzburg, Germany. Her current research interests include the effects of environmental stressors on human health, with specific application to urban environments.

Thilo Erbertseder received the Diploma degree in geography with minors in remote sensing and bioclimatology from the Ludwig-Maximilian University, Munich, Germany, in 1998.

Since 1998, he has been an Atmospheric Scientist and Project Manager with the German Remote Sensing Data Center of DLR, the German Aerospace Center. He has coordinated or participated in more than 50 national and international research projects and has managed large international consortia. His work covers atmospheric composition research, urban climate change, air quality, and health. He is striving to combine global Earth Observation, urban climate modeling, and health risk assessment to make cities livable, healthy, and resilient.

Frank Baier received the Ph.D. degree in realistic stratospheric ozon by assimilation of remote sensing data from the University of Cologne, Institute of Geophysics and Meterolog, Cologne, Germany, in 2000.

He was trained as a Physicist with the University of Cologne, Cologne, Germany. He was active in the German Ozone Research Program OFP and joined the WMO scientific advisory board on ozone. Since 2001, he has worked for the German Aerospace Center as an expert in the analysis of atmospheric trace gases using chemical transport modeling and remote sensing.

Heiko Paeth received the graduated degree in meteorology and geology from the Rheinische Friedrich Wilhelm University Bonn, Bonn, Germany, in 1997, and the doctoral degree in meteorology with subsidiary subject pedology from the University of Bonn, in 2000.

He worked with the Meteorological Institute from 2000 to 2005, and also completed his habilitation there. Since 2006, he has been a Professor of physical geography with the Institute of Geography and Geology, the University of Würzburg, Würzburg, Germany. His work is particularly concerned with climate change, seasonal climate prediction, and climate modeling, with regional focuses on Central Europe, the Mediterranean, and Africa .

Tobias Ullmann studied geography, with majors in physical geography and remote sensing and minors in geology and statistics, at the University of Wrzburg, Wrzburg, Germany, in 2005–2011. He received the Doctorate (Dr. rer. nat.) degree, with topic "Characterisation of Arctic Environment by Means of Polarimetric Synthetic Aperture Radar (PolSAR) Data and Digital Elevation Models (DEM)," from the Graduate School of Science and Technology (GSST), University of Würzburg, Germany, in 2015.

Since 2023, he has been a Full University Professor W2 (Remote Sensing in Geography) at the Department of Remote Sensing of the University of Würzburg. In his scientific work, he deals intensively with the analysis of Earth observation data for current questions of geographical research against the background of climate change. He is an expert in a wide range of remote sensing methods for the digital acquisition, modeling, and analysis of landscape elements and their dynamics. His research interests include the characterization of the Earth's surface and its dynamics.

Hannes Taubenböck received the Diploma degree in geography from the Ludwig-Maximilian University, Munich, Germany, in 2004, and the Ph.D. (Dr.rer.nat.) and the Habilitation degrees in geography from the Julius-Maximilian University of Würzburg (JMU), Würzburg, Germany, in 2008 and 2019, respectively.

He is with the German Aerospace Center (DLR), Weßling, Germany, as well as with the JMU. At the German Remote Sensing Data Center (DFD) of the DLR, he heads the Department "Georisks and Civil Security" and with JMU, he holds the Chair of "Global Urbanization and Remote Sensing". His research interests include remote sensing topics in the domains of urbanization and risk, from the development of algorithms for information extraction to value adding to classification products for findings in urban geography.