

INTEGRATED METHODOLOGICAL FRAMEWORK FOR AIR QUALITY AND HUMAN HEALTH RISK ASSESSMENT IN URBAN AREAS

Reneta Dimitrova¹, Angel M. Dzhambov², Angel Burov^{2,3},
Danail Brezov⁴, Margret Velizarova¹, Veronika Dimitrova⁵,
Nevena Germanova⁶, Elena Hristova⁷

Received on June 5, 2026

Presented by K. Stoykova, Corresponding Member of BAS, on July 28, 2026

Abstract

A multi-scale, multi-model, and data-integrated methodology was developed and applied to assess urban air quality, emission sources, and associated health impacts in Sofia within the framework of the project “Development of a methodology for air quality and human health risk assessment in urban areas”. The project brought together a multidisciplinary team to substitute a shift toward more integrated urban environmental health research in Bulgaria. The methodology is designed to address urban planning needs through numerical scenario analyses, supporting the identification of effective pollution reduction and health risk mitigation measures, while providing a scientific basis for evidence-based decision-making.

Key words: multi-scale modelling approach, data-integrated methodology for urban areas, detailed bottom-up local emissions inventory, air pollution, health risk assessment

1. Introduction. Environmental health research in Bulgaria has often followed a fragmented pattern, with environmental processes and health outcomes examined separately rather than in a unified way. Despite meaningful advances in

We acknowledge funding by the Bulgarian National Science Fund under the project “Development of a methodology for assessing air quality and its impact on human health in urban environments”, grant No. KP-06-H54/2, and the access to the Nestum cluster, HPC Laboratory, Research and Development and Innovation Consortium, Sofia Tech Park.

<https://doi.org/10.7546/CRABS.2026.09.06>

exposure science and in assessing the health effects of air pollution, these areas are still not sufficiently integrated to support contemporary epidemiological studies of long-term residential exposure [1].

Most existing exposure studies rely predominantly on modelling approaches without explicit linkage to health outcomes. Conversely, health studies often depend on coarse exposure indicators, such as data from fixed monitoring stations or proximity-based metrics, which fail to capture fine-scale spatial variability within urban environments. Such limitations restrict the ability to assess intra-urban exposure differences that are critical for robust health impact assessments [1].

A key challenge lies in aligning modelling frameworks with their intended applications. Regulatory and planning-oriented models are typically designed to identify pollution hotspots and ensure compliance, often prioritizing computational efficiency. In contrast, health-focused applications require accurate representation of typical exposure levels at residential locations. This need is reinforced by evidence that adverse health effects may occur even at pollutant concentrations below regulatory thresholds. Spatial resolution is another critical factor, as pollutant concentrations can vary significantly over short distances within urban areas. Conventional coarse-resolution models are generally unable to resolve these gradients, limiting their applicability in epidemiological research.

In response to these challenges, this study develops and applies a multi-scale, multi-model, data-integrated methodology for assessing urban air quality, emission sources and scenarios, and associated health impacts. The approach is demonstrated using the city of Sofia as a pilot case, but it is designed to be transferable to other urban environments.

The methodology integrates high-quality air pollution and meteorological measurements, micro-level health data, disaggregated activity data for emission inventory development, and a modelling framework designed to promote scientific reliability and coherence. It is intended to support urban planning by enabling scenario analysis of mitigation strategies and providing evidence-based inputs for decision-making.

2. Methodology architecture. Detailed methodological architecture was developed to support data collection, including targeted field experiments, the application of novel approaches for addressing data gaps, and the construction of improved emission inventories for the key sources contributing to air pollution. These datasets form the critical input required for the integrated modelling framework applied in air quality assessment and related health analyses. The overall structure, including the main processing steps and data flows of the proposed methodology, is illustrated in Fig. 1.

2.1. Models. A coupled regional-local deterministic modelling framework was implemented for Sofia urban area [2] and employed to capture both large-scale atmospheric processes and fine-scale urban concentration gradients.

At the regional scale, air quality was simulated using the WRF-SMOKE-

A key limitation in Bulgaria is the relatively sparse air quality monitoring network, which constrains LUR model development. Predictor variables typically include traffic-related factors (e.g., distance to roads, traffic intensity), proximity to emission sources, population density, and land-use/land-cover characteristics. Most predictors can be derived from open geospatial datasets (e.g., Copernicus, OpenStreetMap) and are calculated across multiple buffer scales. Traffic variables are most influential at short distances (25–100 m), while land-use variables typically act at broader spatial scales.

While LUR models are computationally efficient and flexible, their performance depends strongly on the availability and quality of input data. To address these limitations, hybrid approaches have been developed that combine LUR with deterministic models or geostatistical methods. These combined frameworks can improve predictive accuracy by integrating physically based simulations with empirical relationships. However, they may also introduce additional uncertainty, especially when relying on input data from previously modelled results. Despite this, such approaches generally require less computational effort than fully deterministic models.

2.2. High-resolution emission inventories. Emission inventories for the transport and domestic heating sectors remain among the most uncertain components of air quality modelling. A central methodological component involves the development of tailored, high-resolution emission inventories, with particular emphasis on transport and residential heating, including the use of solid fuels such as wood and coal, as well as diffuse sources such as bare soil and resuspended dust. To improve the consistency between measured and modelled concentrations, targeted approaches, methods, and tools were applied to better characterise these emission sources. This process required the collection and integration of additional datasets, including detailed information on traffic flow dynamics and fleet composition along major and secondary roads, which enabled the derivation of representative daily and weekly traffic profiles. Complementary datasets include land cover and land use, building volumes, urban morphology, street and junction typologies, and space syntax indicators. Additional inputs comprise estimates of the annual average mileage of registered vehicles at the district level, together with the proportion of total kilometres driven within urban areas. Furthermore, spatial identification of locations where residential heating relies on solid fuels was carried out, supported by data on building typology, residential building census microdata, and heating infrastructure availability.

2.3. Cross-sectional epidemiological study. Robust exposure-response analysis depends on access to reliable and consistent health data. However, due to logistical constraints and the limited duration of the project, the present study relied on self-reported health information. To enhance data reliability, participants were requested to report only physician-diagnosed conditions and corresponding medication use. Despite this approach, self-reported data may still underestimate

disease prevalence, particularly for conditions such as hypertension and diabetes, and their validity can vary across diseases and sociodemographic groups.

Alternative approaches to obtain objective health data in Bulgaria are currently constrained by fragmented and non-standardised health information systems, which limit both data accessibility and interoperability. More resource-intensive study designs, including clinical and case-control approaches, may additionally face constraints related to representativeness and external validity.

Considering these structural limitations, a practical direction for future research, is the systematic enhancement of existing national health surveys (e.g. the European Health Interview Survey) through the integration of additional modules addressing air pollution exposure, residential history, and time-activity patterns. This approach may be further strengthened by the inclusion of geospatial information, enabling the assessment of co-exposures such as environmental noise and access to green space. Targeted validation of selected subsamples through linkage with medical records or clinical assessments could also be considered, thereby improving confidence in exposure-response relationships and supporting more robust evidence for policy development.

2.4. Field experiments. Field experiments represent an important methodological component for improving the understanding of air pollution processes across multiple spatial scales (national, urban, and street level). They provide direct observations under real-world conditions, enabling a more reliable assessment of interactions between emission sources, atmospheric processes, and meteorological factors. The generated data support the development and validation of air quality models and provide a foundation for informing environmental standards, regulatory policies, and mitigation measures. In their absence, analyses may rely on simplified assumptions, potentially leading to less reliable conclusions. Field measurements therefore enhance understanding of pollution gradients and support more robust, evidence-based decision-making.

3. Results and discussion. The developed methodology was applied to the pilot area of Sofia, the largest city in the country. Despite sustained efforts by both national and local authorities, its population continues to be exposed to elevated levels of particulate matter (PM) and nitrogen dioxide (NO₂).

Sofia constitutes a complex and challenging study area due to its geographical setting and urban characteristics. The city is located in a basin with complex terrain and meteorological conditions that favour pollutant accumulation. These natural constraints are further exacerbated by rapid urban expansion and recent urban development trends, including high-density construction and the reduction of inter-block green spaces. Additional factors negatively affecting air quality include the increasing cover of impermeable surfaces, the formation of urban street canyons, disruption of ventilation corridors such as “green wedges”, and transport infrastructure developments that encourage private vehicle use.

3.1. Database construction and emission inventory development.

Data availability for key domains, such as transport, residential heating, and urban green spaces remains limited and subject to significant temporal variability. An initial assessment of datasets related to meteorology, air quality, population health, transport, residential heating, and urban development identified substantial gaps, inconsistencies, and misalignments in spatial and temporal resolution. To mitigate these limitations, multiple data sources and access platforms were systematically explored. A structured data harmonisation and integration procedure was subsequently implemented within the methodological framework and parametric software environment, with the aim of improving representativeness while maintaining resource efficiency.

Targeted field measurements of PM were conducted at three representative site types: a high-traffic urban boulevard, an urban park, and a background site located outside the urban area. The scarcity of such in situ observations in Sofia highlights the added value of the newly collected dataset, which enables a more robust assessment of spatial pollution gradients under conditions of marked surface heterogeneity, thereby supporting evidence-based policy and planning decisions.

Satellite data from Sentinel-2 imagery were also collected and applied to the estimation of land-cover related indices. The resulting dataset supports the assessment of the influence of urban morphology and land cover on pollutant dispersion and deposition processes. These insights contribute to a more accurate representation of local-scale dynamics in air quality modelling and provide relevant evidence to inform urban planning and land-use management strategies.

A diverse and heterogeneous traffic dataset was compiled and systematically processed, covering major boulevards, key intersections, and a range of street typologies across the urban area. Data sources comprised machine-readable measurements from the Road Infrastructure Agency; datasets from the municipal Directorate “Traffic Management and Analysis” (via Sofiaplan), covering selected periods between 2017 and 2019 at approximately 40 junctions; targeted traffic surveys conducted by consultancy firms and non-governmental organizations; and outputs from the Open Transport Model (2015).

Data processing, analysis, and visualisation were carried out using QGIS v.3.22-40 and Python-based tools. The methodological approach incorporated an ensemble learning framework, integrating a broad set of explanatory variables, including road classification, population density, functional land use characteristics, space syntax indicators, as well as traffic observations and model-derived outputs.

A set of advanced regression techniques, namely Random Forest, XGBoost, and CatBoost, were applied. Model performance was assessed using standard statistical metrics, and the individual models were subsequently combined into a weighted ensemble to improve predictive performance. Further methodological details are provided in [9, 10].

The resulting activity data included a detailed fleet characterisation comprising 546 vehicle subcategories, defined by vehicle type (motorcycles, light- and heavy-duty vehicles), fuel type and propulsion technology, emission standards (Euro classes), vehicle mass, and emission control conditions. These were combined with additional parameters such as traffic intensity, speed profiles, road surface characteristics, gradients, and street canyon morphology. The latter were obtained using the Urban Canyon Tool integrated into ADMS-Urban and implemented within the ArcGIS 10.6.1 environment.

The integrated dataset was subsequently used as input to the EMIT module to estimate transport-related emissions across the primary road network. The resulting high-resolution emission inventory, generated at the level of individual road segments, was then applied as input to the ADMS-Urban dispersion model to support air quality simulations and scenario-based policy assessment.

3.2. Outcomes of the applied methodology. The ADMS-Urban model was applied to estimate current concentrations of key air pollutants. Prior to its implementation, an extensive validation process was undertaken [2], including the refinement of emission factors for non-exhaust PM sources and adjustment of NO_x fraction relationships based on empirical functions derived from measurements conducted in Sofia [11]. Meteorological conditions, road surface characteristics, and traffic dynamics in Bulgaria differ substantially from those in the United Kingdom and other European countries, where many default modelling assumptions originate. To account for these differences, locally developed methodologies were applied to observational datasets to derive site-specific parameterisations, as described in [11]. This tailored approach improved agreement between modelled and observed pollutant concentrations, thereby enhancing the reliability of the modelling outputs for policy and planning applications.

Additional analysis focused on PM resuspension; a source category associated with considerable uncertainty due to its strong seasonal variability and sensitivity to dry conditions. To better quantify this contribution, an integrated dataset was developed that combines remote sensing indices – Normalized Difference Vegetation Index (NDVI), Bare Soil Index (BSI), and Normalized Difference Moisture Index (NDMI), derived from Sentinel-2 imagery, with meteorological and air quality observations. These data were further complemented by variables describing traffic intensity, emission inventory estimates, transport infrastructure, and demographic characteristics, including the spatial distribution of the motorised population and the prevalence of residential heating using wood and coal. Statistical and machine learning methods were applied to evaluate the contribution of surface conditions and bare soil to ambient PM concentrations, as described in [12]. This approach improved the representation of non-exhaust emission processes and their spatial variability, thereby supporting more targeted air quality management and policy interventions.

A coupled regional-local modelling framework was implemented to assess the effectiveness of low-emission zones (LEZs) introduced in central Sofia [13]. Emission inventories were developed for three scenarios: a baseline year (2018), a small-ring LEZ scenario (2022), and a large-ring LEZ scenario (2026). The results indicate reductions in NO₂ concentrations along major roadways and at high-traffic intersections under LEZ scenarios. However, the projected concentrations remain above the stricter limits set by the forthcoming EU Air Quality Directive, effective from 2030. In contrast, PM levels show only limited improvement, highlighting the persistent influence of non-exhaust emissions and other sources. Further details are provided in [13].

Finally, the relationships between greenspace, traffic-related air pollution, environmental noise, and self-reported health outcomes were analysed using data from the cross-sectional epidemiological study [14]. To ensure sufficient variability in exposure, a clustered sampling design was applied, defining eight spatial typology classes based on combinations of high and low levels of greenspace, air pollution, and traffic noise.

The analysis, described in detail in [14], shows that residence in greener neighbourhoods and access to private gardens are associated with a lower prevalence of poor self-rated health. In contrast, higher exposure to traffic-related air pollution, as well as aircraft and railway noise, is linked to adverse health outcomes. These findings underscore the need for integrated urban planning approaches and highlight the importance of coordinated policies addressing transport, air quality, noise, and public health. Furthermore, the results emphasise the role of socio-economic factors, suggesting that mitigation measures should prioritise vulnerable population groups, including individuals with lower educational attainment and limited financial resources.

4. Conclusions. The proposed methodology constitutes an integrated regional-local modelling framework that combines high-resolution, locally adapted emission inventories, advanced machine learning techniques for data fusion and gap filling, targeted field measurements, and a cross-sectional epidemiological study. This multi-component approach provides a structured basis for interdisciplinary research on the monitoring and modelling of meteorological conditions, air quality, and associated health impacts.

The application of these methods supports the development and evaluation of scenario-based analyses, facilitating the identification of effective emission reduction strategies and health-oriented interventions. In addition, the framework enables the generation of evidence-based inputs for decision-making, contributing to more informed policy development and the advancement of sustainable urban planning.

REFERENCES

- [1] DZHAMBBOV A. M., D. DIMITROVA, A. BUROV et al. (2025) Epidemiology of Air Pollution and Health, Plovdiv, Avto Print, ISBN: 978-619-7205-40-4, (in Bulgarian).
- [2] DIMITROVA R., M. VELIZAROVA, G. GADZHEV et al. (2025) Implementation of coupled regional-to-local scale modelling system in Bulgaria – air quality study for the city of Sofia, C. R. Acad. Bulg. Sci., **78**(2), 393–401.
- [3] SKAMAROCK, W. C., J. B. KLEMP, J. DUDHIA et al. (2008) A description of the advanced research WRF Version 3, NCAR – National Center for Atmospheric Research, Boulder, CO, USA.
- [4] Institute for the Environment, The University of North Carolina at Chapel Hill (2017) SMOKE v4.5 User's Manual, CMAS Center.
- [5] United States Environmental Protection Agency (2012) CMAQ (Version 5.0) [Software], Available online: <https://zenodo.org/records/1079888>, (accessed on 20 May 2026).
- [6] Cambridge Environmental Research Consultants Ltd., Urban Air Quality Management System, Version 5.0, User Guide, Available online: <https://www.cerc.co.uk/environmental-software.html>, (accessed on 20 May 2026).
- [7] Cambridge Environmental Research Consultants Ltd., Multi-Model Air Quality System, Version 1.2, User Guide, Available online: <https://www.cerc.co.uk/environmental-software.html>, (accessed on 20 May 2026).
- [8] Cambridge Environmental Research Consultants Ltd., Atmospheric Emissions Inventory Toolkit (EMIT), Version 3.9, User Guide, Available online: <https://www.cerc.co.uk/environmental-software.html>, (accessed on 20 May 2026).
- [9] BUROV A., D. BREZOV (2023) Transport Emissions from Sofia's Streets – Inventory, Scenarios, and Exposure Setting. In: EnviroRISKs 2022, LNNS 638 (Eds N. Dobrinkova and O. Nikolov), 223–233.
- [10] BREZOV D., A. BUROV (2023) Ensemble learning traffic model for Sofia: A case study, Appl. Sci., **13**(8), 4678.
- [11] VELIZAROVA M., R. DIMITROVA, P. O. HRISTOV et al. (2024) Evaluation of emission factors for particulate matter and NO₂ from road transport in Sofia, Bulgaria, Atmosphere, **15**(7), 773.
- [12] BREZOV D., R. DIMITROVA, A. BUROV et al. (2025) Analyzing the contribution of bare soil surfaces to resuspended particulate matter in urban areas via machine learning, Appl. Sci., **15**(23), 12783.
- [13] DIMITROVA R., M. VELIZAROVA, A. BUROV et al. (2025) Numerical simulations and assessment of the effect of low-emission zones in Sofia, Bulgaria, Urban Sci., **9**(10), 402.
- [14] DZHAMBBOV A. M., V. DIMITROVA, N. GERMANOVA et al. (2023) Joint associations and pathways from greenspace, traffic-related air pollution, and noise to poor self-rated general health: A population-based study in Sofia, Bulgaria, Environ. Res., **231**(1), 116087.

¹Department of Meteorology and Geophysics, Faculty of Physics,
Sofia University “St. Kliment Ohridski”, 5 James Bourchier Blvd, 1164 Sofia, Bulgaria
e-mails: r.dimitrova@phys.uni-sofia.bg,
<https://orcid.org/0000-0002-5931-8713>
margret@phys.uni-sofia.bg

²Environmental Health Division, Research Institute at Medical University of Plovdiv,
Medical University of Plovdiv, 15A Vasil Aprilov Blvd, 4002 Plovdiv, Bulgaria
e-mail: angel.dzhambov@mu-plovdiv.bg, <https://orcid.org/0000-0003-2540-5111>

³Department of Urban Planning, Faculty of Architecture, University of Architecture,
Civil Engineering and Geodesy, 1 Hristo Smirnenski Blvd, 1046 Sofia, Bulgaria
e-mail: Burov_far@uacg.bg, <https://orcid.org/0000-0002-8631-5462>

⁴Department Mathematics, Faculty of Transportation Engineering,
University of Architecture Civil Engineering and Geodesy, 1 Hristo Smirnenski Blvd,
1046 Sofia, Bulgaria
e-mail: danail.brezov@gmail.com, <https://orcid.org/0000-0003-4157-1072>

⁵Department of Sociology, Faculty of Philosophy, Sofia University “St. Kliment Ohridski”,
15 Tsar Osvoboditel Blvd, 1504 Sofia, Bulgaria
e-mail: v.s.dimitrova@phls.uni-sofia.bg/veronika.s.dimitrova@gmail.com
<https://orcid.org/0000-0001-8116-0863>

⁶Department of Spatial and Strategic Planning of Sofia Municipality, Sofiaplan, Zone B5,
Building 11, Entr. B, Floor 17, Sofia, Bulgaria
e-mail: nevena@sofiaplan.bg

⁷Department of Meteorology, National Institute of Meteorology and Hydrology,
66 Tsarigradsko Shosse Blvd, 1784 Sofia, Bulgaria
e-mail: elena.hristova@meteo.bg, <https://orcid.org/0000-0002-6159-3554>