Dynamic exposure modeling and personalized risk assessment in urban environments: a data-driven approach based on heterogeneous IoT networks*

Kyrylo Vadurin[1,*,†], Andrii Perekrest[1,†], Volodymyr Bakharev[1,†] and Dmytro Mamchur[1,†]

[1] *Kremenchuk Mykhailo Ostrohradskyi National University, Universytetska str., 20, 39600 Kremenchuk, Ukraine*

Abstract

This study addresses the discrepancy between static municipal air quality monitoring and actual physiological exposure in dynamic urban environments. We propose a hybrid framework for personalized risk assessment that integrates high-resolution heterogeneous data from the EcoCity Internet of Things network with agent-based simulation and physiological ventilation models. The research formalizes an integral risk index, which accounts for real-time changes in pulmonary ventilation and micro-environmental context (indoor/outdoor transitions). The methodology was validated using a longitudinal dataset of 1.4 million records from the Vinnytsia agglomeration (2019–2025). The results demonstrate that traditional static monitoring fails to detect critical health risk events in 47.6% of simulated high-intensity activity scenarios for Vulnerable users. Furthermore, a segmented context-aware route optimization algorithm is presented, achieving a measurable reduction in the accumulated inhalation dose. This work provides a scalable decision support tool for smart cities, shifting the paradigm from environment-centric observation to human-centric health protection.

Keywords

air quality monitoring, IoT sensor networks, personalized risk assessment, digital twins, PM2.5, human-centric simulation, decision support systems, agent-based modeling

1. Introduction

1.1. Urban environmental challenges and digital transformation

The rapid pace of global urbanization and industrial development has precipitated a complex array of environmental challenges that directly impact human health and the sustainability of urban ecosystems. As cities densify, the interplay between anthropogenic emissions, meteorological dynamics, and urban topology creates highly heterogeneous pollution fields that are difficult to monitor using traditional static infrastructure. The paradigm of smart cities has emerged as a response to these challenges, advocating for the integration of digital technologies to enhance the management of urban resources and quality of life. Within this context, the monitoring of atmospheric air quality has shifted from a peripheral regulatory obligation to a central component of public health protection strategies. However, the sheer volume and velocity of data generated by modern observation networks necessitate the transition from simple monitoring to complex forecasting and decision support systems, a necessity highlighted in recent studies on digitalization for air pollution detection [1].

Recent literature emphasizes that sustainable development goals cannot be achieved without a rigorous, data-driven approach to environmental management. The concept of "Precision urban health" is emerging as a critical discipline, requiring the ability to tailor environmental

---

[*] *CMIS-2026: The Ninth International Workshop on Computer Modeling and Intelligent Systems, May 05, 2026, Zaporizhzhia, Ukraine*

[1*] Corresponding author.

[†] These authors contributed equally.

kir3337@gmail.com (K. Vadurin); pksg13@gmail.com (A. Perekrest); v.s.baharev@gmail.com (V. Bakharev); dgmamchur@gmail.com (D. Mamchur)

0000-0001-7781-5783 (K. Vadurin); 0000-0002-7728-9020 (A. Perekrest); 0000-0001-9312-654X (V. Bakharev); 0000-0002-2851-878X (D. Mamchur)

CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073

interventions and recommendations to specific microclimates and individual physiological vulnerabilities. Research indicates that decision support systems are becoming pivotal in bridging the gap between raw data collection and actionable management strategies [2, 3]. The common denominator in these diverse fields is the necessity to process heterogeneous data streams — including satellite imagery, ground-based sensor readings, and meteorological forecasts — to derive insights that are temporally relevant and physiologically precise [4, 5].

1.2. Evolution of environmental monitoring from static observation to dynamic data fusion

Historically, environmental monitoring has relied on sparse networks of high-precision reference stations. While these stations provide regulatory-grade data, their spatial coverage is insufficient to capture the fine-grained variability of air pollution in complex urban terrains. The limitation of this traditional approach creates a phenomenon often referred to as the information gap, where significant localized pollution events remain undetected, leading to an underestimation of public exposure. To address this, the scientific community has increasingly turned to integrating remote sensing technologies and the Internet of Things (IoT). Remote sensing, particularly through satellite observations, offers the advantage of broad spatial coverage, allowing for the holistic assessment of environmental parameters over large territories. Studies have demonstrated the utility of remote sensing in managing river systems, coastal environments, and agricultural lands, proving its capability to detect trends that ground-based systems might miss [6, 7].

However, remote sensing data often suffers from low temporal resolution and atmospheric interference. Conversely, the IoT enables the deployment of low-cost sensors that provide high temporal resolution but may lack the measurement accuracy of reference stations [8]. The convergence of these technologies through data fusion techniques represents a significant advancement in environmental informatics. By combining the broad spatial coverage of satellite data with the high temporal fidelity of ground-based sensors, researchers can reconstruct continuous spatiotemporal fields of pollutant concentrations. This hybrid approach allows for a more granular understanding of environmental dynamics, yet it introduces new challenges regarding data heterogeneity, noise reduction, and the computational complexity of real-time processing [9].

Furthermore, the integration of these diverse data sources requires sophisticated geospatial modeling. Geographic information systems have evolved from simple mapping tools into powerful analytical engines capable of complex spatial interpolation and geostatistical analysis. Recent work in land suitability assessment and river management demonstrates how geographic information systems can be coupled with multi-criteria decision-making frameworks to optimize the spatial allocation of resources and identify critical areas for intervention [10, 11]. Applying similar principles to air quality monitoring allows for the development of dynamic maps that reflect the real-time dispersion of pollutants, thereby providing a foundation for more responsive environmental management systems.

1.3. The role of geospatial modeling and agent-based simulation

While artificial intelligence and deep learning have become popular tools in environmental forecasting, they often function as "black boxes," lacking the interpretability required for public health decision-making. Consequently, there is a growing need for transparent, explainable modeling approaches that can simulate cause-and-effect relationships in urban environments. Instead of purely stochastic predictions, modern research is pivoting towards the concept of Digital twins and Agent-based modeling (ABM). These methodologies allow for the reconstruction of continuous spatiotemporal fields from discrete data points and the simulation of individual human interactions with these fields [12].

A critical area of development involves the hybridization of data-driven interpolation with physiological models. By integrating geospatial analysis with biological ventilation parameters, researchers can estimate not just the state of the atmosphere, but the actual inhalation dose received by an individual. This approach moves beyond the capabilities of classical autoregressive models, which typically treat the population as stationary receptors. The application of such hybrid frameworks to urban environmental forecasting holds the promise of significantly reducing "false negative" errors in risk assessment, particularly for volatile parameters such as particulate matter concentrations during peak activity periods. The ultimate goal is to create intelligent systems that provide physically consistent and explainable recommendations to decision-makers, thereby fostering trust and facilitating the adoption of automated tools in municipal governance [16].

1.4. Limitations of current exposure assessment methodologies

Despite the technological advancements in data collection and forecasting, a critical disconnect remains between the measurement of environmental parameters and assessing their actual impact on human health. Traditional environmental risk assessment relies heavily on static metrics, such as the time-weighted average of pollutant concentrations over fixed periods, for example, twenty-four hours or one year. This approach assumes that human exposure is uniform over time and space, effectively treating the population as stationary receptors. However, human mobility patterns and physiological variations significantly influence the actual inhaled dose of pollutants.

Research in behavioral geography and exposure science indicates that individuals constantly move between different micro-environments — homes, offices, transit corridors, and outdoor recreational areas — each characterized by distinct pollution levels. Furthermore, the intensity of physical activity dramatically alters the ventilation rate, determining the volume of air—and consequently the mass of pollutants — inhaled per unit of time. A person waiting for a bus along a busy street inhales a significantly different dose than someone sitting in an air-conditioned office, even if the outdoor air quality reading for the district is identical. The failure to account for these dynamic factors leads to a systematic error in risk assessment, often referred to as the false negative error of static monitoring, where hazardous exposure events during peak activity are masked by lower average concentrations [1].

Current decision support systems for environmental management often lack the functionality to model these dynamic exposure scenarios. While systems exist for optimizing wildfire suppression, managing agricultural supply chains, or planning urban infrastructure, few integrate the physiological dimension of human exposure into their logic [5, 8]. Most existing solutions focus on the environment-centric view — optimizing sensor placement or predicting concentration fields — rather than the human-centric view of minimizing the received toxic dose. This limitation is particularly critical for vulnerable population groups, such as individuals with asthma or cardiovascular conditions, for whom short-term exposure to high pollution peaks during physical exertion can trigger acute health crises [9].

1.5. The need for systemic thinking and integrated decision support

The transition from monitoring environmental parameters to actively managing health risks requires a systemic approach to decision-making. Systemic thinking involves understanding the interdependencies between various components of the urban system: the physical environment, the built infrastructure, human behavior, and the regulatory framework [4]. Applications of systemic thinking in decision-making have highlighted the importance of identifying synergies and trade-offs between different sustainable development goals. For example, a policy promoting active transportation to reduce carbon emissions might inadvertently increase the population's exposure to traffic-related air pollution if not accompanied by appropriate route planning and infrastructure design.

Developing decision support systems that facilitate such systemic thinking is a complex engineering challenge. It requires the integration of heterogeneous modules: data acquisition pipelines, predictive models, simulation engines, and user interfaces for risk communication. Moreover, these systems must be adaptive, capable of learning from new data and feedback to refine their recommendations [17]. The concept of the digital twin offers a promising architectural pattern for such systems. By creating a virtual replica of the urban environment and the agents within it, researchers can simulate various scenarios — such as the impact of a new traffic regulation or the health consequences of a specific route choice — before implementing them in the real world [18].

The design of such systems must also consider the user engagement aspect. A technically sophisticated tool is of little value if its outputs are not interpretable or actionable for the end-user. Studies on user engagement with decision support tools emphasize the need for intuitive visualizations, such as color-coded risk maps and clear, context-aware alerts [19]. The effectiveness of a decision support system is measured not just by the accuracy of its algorithms, but by its ability to influence behavioral change and improve decision-making outcomes. In the context of air pollution, this means empowering citizens to make informed choices about their daily activities and mobility patterns to minimize their personal health risks.

1.6. Research problem and objectives

The overarching problem addressed in this study is the inadequacy of existing static monitoring frameworks to protect public health in dynamic urban environments. While significant progress has been made in the digitalization of environmental monitoring, current systems predominantly function as data archives rather than proactive risk management tools [1]. They fail to bridge the gap between the measured concentration of pollutants and the actual biological dose received by individuals, particularly during periods of physical activity. Furthermore, existing navigation and routing services optimize primarily for distance or time, ignoring the cumulative toxic load associated with different trajectories through the urban pollution field.

This research aims to develop a comprehensive methodology for dynamic exposure modeling and personalized risk assessment within a smart city context. The study builds on previous work in environmental information systems, expanding the scope to include agent-based simulation and physiological modeling [14]. The specific objectives of this work are multifarious.

First, to develop a hybrid data fusion methodology that integrates historical data from stationary monitoring posts with geospatial interpolation techniques and context-aware coefficients. This aims to reconstruct high-resolution spatiotemporal pollution fields that account for the variation between indoor and outdoor environments, thereby reducing the uncertainty associated with sparse sensor networks [20].

Second, to formulate a mathematical model for the integral risk index that dynamically calculates the accumulated inhaled dose of pollutants based on real-time pollution levels and the variable ventilation rates associated with different human activities. This model seeks to provide a more physiologically relevant metric for health risk assessment than the standard air quality index.

Third, to design and validate a Segmented context-aware detour algorithm for spatial optimization of pedestrian routes. Unlike traditional pathfinding algorithms, this approach prioritizes the minimization of the accumulated toxic dose while maintaining the logical structure of the user’s daily itinerary, such as the necessity to visit specific waypoints [16].

Fourth, to demonstrate the efficacy of the proposed system through a comparative analysis of static versus dynamic monitoring approaches using historical data from the industrial agglomeration of Vinnytsia, Ukraine. This validation intends to quantify the information gap and provide statistical evidence of the system's ability to detect critical risk events that are missed by conventional methods.

1.7. Scientific novelty and contribution

The scientific novelty of this research lies in the further development of the method for dynamic assessment of aerogenic risk in urbanized environments. Unlike existing approaches based on time-weighted average indicators, the proposed method integrates the variable parameter of pulmonary ventilation into the model of dose accumulation in real time. This allows for the formalization and quantitative estimation of the discrepancy between data from stationary monitoring posts and the actual exposure of a mobile user, significantly reducing the probability of type two errors when detecting critical states for vulnerable population groups [1, 13].

Furthermore, the method of spatial optimization of pedestrian routes has been improved through the application of segmented vector heuristics in a continuous pollution field. This modification enables the identification of compromise trajectories that preserve mandatory intermediate route points while ensuring a reduction in the calculated inhalation dose, which distinguishes it from classical shortest-path algorithms [16].

Finally, the approach to aggregating heterogeneous environmental data has been further developed. The hybrid model combines direct measurements, geospatial interpolation, and contextual coefficients of pollution penetration into premises. This integration enables a holistic analysis of the environment-human system, moving beyond the isolated monitoring of atmospheric parameters to a comprehensive assessment of environmental safety and quality of life [9].

1.8. Structure of the study

The subsequent sections of this paper are structured to provide a logical progression from theoretical foundations to practical validation. The methodology section details the data sources, preprocessing techniques, and mathematical formulations used for geospatial modeling and risk calculation. The results section presents the outcomes of the spatial pollution analysis, dynamic exposure simulation, and the comparative validation of the risk assessment model. The decision support system section describes the logic of the risk management alerts and the route optimization algorithm. Finally, the discussion and conclusion sections interpret the findings in the broader context of urban environmental management, acknowledge the study's limitations, and outline directions for future research [5]. Through this structure, the paper demonstrates how the integration of advanced computational methods can transform environmental monitoring from a passive observational science into an active tool for health protection.

2. Materials and methods

2.1. Data acquisition and preprocessing pipeline

The empirical foundation of this study rests upon a dual-source data strategy. The primary source of environmental data was the EcoCity public monitoring network, specifically the archival dataset "Air Quality Monitoring from EcoCity" for the Vinnytsia agglomeration. This dataset is characterized by a high temporal resolution (15 seconds to several minutes), spanning from February 2019 to March 2025, providing approximately 1.4 million data points. To complement the physical data, the study utilized a synthetic "Air Quality and Health Impact Dataset" to demonstrate the logic of the risk categorization module. It is important to note that this synthetic dataset serves solely for the calibration of the system's internal classifiers and the demonstration of the decision support logic, rather than for deriving clinical medical conclusions. The integration of these sources addresses the data availability challenges often cited in environmental research.

The data preprocessing pipeline involved a rigorous cleaning and synchronization phase. Raw sensor data were filtered to remove artifacts resulting from sensor drift or connectivity interruptions. Following outlier removal, time series from distributed stations were resampled into consistent 20-minute intervals using linear interpolation. This temporal standardization was essential to enable the application of the RegularGridInterpolator for field reconstruction. The preprocessing methodology ensures that the input vectors for the risk model maintain statistical integrity [14].

2.2. Geospatial modeling and continuous field reconstruction

To transition from discrete point measurements provided by the monitoring stations to a continuous representation of the urban environmental quality, a geospatial modeling approach was employed. Traditional monitoring networks suffer from spatial sparsity, leaving vast areas of the urban landscape unmonitored—a phenomenon that creates significant information gaps in risk assessment. To bridge this gap, the study utilized a spatial interpolation technique based on the principle of spatial decay rather than simple linear triangulation. This method posits that the influence of a pollution source detected by a sensor diminishes as the distance from the source increases, following a non-linear decay function governed by atmospheric dispersion physics.

The mathematical formulation of the pollution field was implemented using a grid-based approach where the study area was discretized into a uniform high-resolution mesh. The value of each cell in this grid represents the estimated PM2.5 concentration at that specific geospatial coordinate. The estimation algorithm integrates readings from all active stations within a defined radius of influence (r_{max}), weighted by their distance to the grid cell. The influence intensity (I) of a sensor at a given point is calculated using a distance-decay function:

$$I = V_{sensor} \cdot \left(1 - \left(\frac{d}{r_{max}}\right)^{p}\right), \tag{1}$$

where V_{sensor} is the measured value at the station, d is the Euclidean distance between the grid cell and the station, and p is the decay power coefficient, which determines how rapidly the influence fades. In this study, a decay power of $p = 1.5$ and a maximum radius of 8000 meters were empirically selected to reflect the dispersion characteristics of fine particulate matter in a semi-urban terrain such as Vinnytsia. This approach ensures that the model accounts for the "hotspot" nature of urban pollution, where local sources create peaks that override the background pollution levels [7].

The geospatial module also incorporates contextual data from OpenStreetMap (OSM) to provide semantic meaning to the coordinate grid. By overlaying the calculated pollution field onto the urban topology—including road networks, building footprints, and green zones — the system can contextually analyze exposure risks. For instance, the model distinguishes between open spaces where dispersion is rapid and street canyons where pollutants may accumulate. The visualization of these fields uses a sequential color ramp (e.g., from green to red or white to purple) to intuitively communicate the intensity of pollution to the user. This integration of remote sensing principles with GIS technologies creates a comprehensive digital substrate for the subsequent agent-based simulation, advancing the methods described in [6] for environmental management systems. The resulting scalar field provides a dynamic heatmap of air quality, allowing for the identification of safe and hazardous zones within the city at any given timestamp.

2.3. Mathematical formulation of the integral risk index

The core theoretical contribution is the transition from a concentration-based to a dose-based risk assessment. The Integral risk index R_{int} accumulates the instantaneous inhaled dose over the entire exposure period:

$$R_{\text{int}} = \sum_{t=0}^{T} \left(C(x_t, y_t) \cdot V_{rate}(Activity_t) \cdot K_{tox} \cdot K_{sens} \right), \tag{2}$$

where $C(x_t, y_t)$ is the ambient concentration at the agent's location, and V_{rate} is the respiratory minute volume. The coefficient K_{tox} represents the inherent toxicity of the specific pollutant. In this baseline model, K_{tox} is set to 1.0 for PM2.5. To ensure the integral nature of the risk index and its full extensibility for multi-pollutant environments, K_{tox} acts as a toxicity multiplier based on relative health impact guidelines (e.g., WHO standards). For instance, when integrating other pollutants into the total risk score, K_{tox} can be explicitly defined as 1.2 for Nitrogen Dioxide and 0.8 for Sulfur Dioxide to reflect their relative physiological impact potentials compared to fine particulate matter. K_{sens} represents individual susceptibility (1.0 for healthy, 1.5 for vulnerable profiles).

Furthermore, the model accounts for the protective effects of indoor environments using an indoor/outdoor (I/O) ratio coefficient. Based on average infiltration rates for naturally ventilated buildings in urban areas, a conservative I/O ratio of 0.6 was adopted for this study [20]. This parameter allows the model to simulate the attenuation of exposure during indoor activities, providing a more realistic accumulation curve than outdoor-only models.

2.4. Agent-based simulation and digital twins

To validate the proposed risk model and explore the dynamics of exposure in a controlled yet realistic setting, an agent-based simulation environment was developed. This approach utilizes the concept of "Digital twins" — virtual representations of physical entities that can be used to simulate complex interactions and predict future states [18]. In this study, two distinct digital twins were instantiated: a "Healthy user" and a "Vulnerable user (Asthma profile)," each governed by the physiological parameters described in the previous section. The use of digital twins allows for the execution of "what-if" scenarios that would be ethically or practically impossible to conduct with human subjects in hazardous environments.

The simulation environment reconstructs a typical daily routine comprising three distinct activity segments: Commuting (walking), Office work (sedentary), and Sport (running). These activities were chosen to represent a broad range of ventilation rates and spatial contexts. The "Commute" segment simulates movement through the urban grid, exposing the agent to varying outdoor pollution levels derived from the interpolated field. The "Office work" segment simulates a stationary period in an indoor environment, testing the model's handling of the I/O ratio and risk stabilization. The "Sport" segment simulates high-intensity activity in an outdoor setting (e.g., a park), representing the critical scenario where high ventilation rates might lead to rapid dose accumulation. The duration of each segment was set to represent a realistic part of a daily schedule, with the sport segment specifically targeted to stress-test the risk accumulation model.

The trajectory of the agents is generated procedurally to ensure spatial realism and relevance to the specific pollution conditions of the simulation date. Instead of using fixed, hard-coded coordinates, the system identifies active monitoring stations with the highest pollution readings. It generates "Home," "Work," and "Park" locations within a randomized radius (typically 1.5 to 2.5 km) around these sensors. This procedural generation ensures that the simulation scenarios focus on the most problematic areas of the city, thereby providing a conservative estimate of the maximum potential risk. The movement between these points is simulated using linear interpolation in the current iteration, with provisions for future integration of graph-based street routing.

The simulation engine computes the state of each agent at 1-minute intervals. At each time step, the system queries the interpolated pollution map for the agent's current coordinates, applies the relevant I/O ratio based on the location type, determines the ventilation rate based on the current activity, and calculates the incremental increase in the I_{risk} index. To introduce stochastic realism and simulate sensor noise inherent in wearable devices, a Gaussian noise component is added to the "sensed" values. This granular, time-step simulation allows for the detection of critical threshold crossings — moments when the accumulated risk exceeds safe limits defined by the World Health Organization guidelines — which would be invisible to temporal aggregation methods like Time-weighted averages (TWA). This methodology aligns with the scenario-based decision analysis approaches described in [17], providing a structured way to evaluate outcomes under uncertainty.

2.5. Algorithm for segmented context-aware routing

A key component of the proposed decision support system is the ability to actively mitigate risk through route optimization. Traditional routing algorithms, such as Dijkstra's algorithm or A*, typically optimize for the shortest distance or fastest time. In the context of environmental health, the objective function must be altered to minimize the accumulated inhalation dose. To achieve this, the study developed a "Segmented context-aware detour" algorithm. Unlike standard pathfinding which treats the route as a single continuous line, this algorithm recognizes the segmented nature of human mobility, where intermediate waypoints (such as the workplace or a specific shop) are mandatory constraints that cannot be skipped.

The algorithm divides the total itinerary into logical segments (e.g., Home to Work, Work to Park) and optimizes each segment independently before concatenating them into a final trajectory. The optimization logic employs a vector-based heuristic rather than a graph-search approach, which is computationally efficient for continuous fields. For a given segment defined by a start point A and an end point B, the algorithm calculates the direct vector and its normal vector. It then generates a candidate "detour" path by displacing the midpoint of the segment along the normal vector, creating a smooth Bezier-like curve. The magnitude and direction of this displacement, controlled by a "bend factor," are iteratively adjusted to probe the pollution field for areas of lower concentration.

For each candidate trajectory, the algorithm calculates the total accumulated dose by integrating the pollution values along the path. This heuristic approach allows the system to find "compromise" routes that deviate slightly from the straight line to avoid pollution hotspots, without requiring the computational overhead of traversing a dense street graph. The "green route" is defined as the trajectory that offers the maximum reduction in accumulated dose within a reasonable constraint of additional distance. This method aligns with the principles of intelligent decision support systems, where machine learning and optimization algorithms are combined to solve multi-criteria problems [14, 16]. The algorithm also implicitly accounts for the "cost" of deviation; if a cleaner path requires an excessively long detour, the increased duration of exposure might negate the benefit of lower concentrations. The integral risk model naturally captures this trade-off, as I_{risk} is a function of both concentration and time.

2.6. Validation and comparative analysis framework

To ensure the scientific rigor of the proposed methodology, a validation framework was established to compare the performance of the dynamic risk model against traditional static monitoring methods. This comparative analysis serves as the A/B testing phase of the research, essential for demonstrating the added value of the new approach and quantifying the limitations of current practices. The validation strategy involves the simulation of a "Municipal station" data stream alongside the "Personal sensor" stream.

The Municipal station stream is generated by averaging the outdoor pollution levels across the entire simulation period and adding random noise to mimic measurement fluctuations typical of stationary sensors. This represents the data available to a city manager or a citizen using a standard weather app — a static, city-wide average that does not account for user location or activity. In contrast, the Personal sensor stream reflects the dynamic, high-resolution data experienced by the agent in the simulation, accounting for local variations in the pollution field and the agent's movement through it.

By plotting the I_{risk} accumulation curves derived from both data streams, the study aims to quantify the "Information gap"—the discrepancy between the perceived risk based on static station data and the actual risk based on dynamic exposure. This gap represents the potential for Type II errors (false negatives), where a user might engage in hazardous activities because the general city air quality is reported as "Moderate," unaware that their specific location and high ventilation rate are pushing their personal toxic dose into the "Critical" zone. The visualization of this gap provides a compelling argument for the adoption of personalized monitoring systems.

Furthermore, the study leverages the historical depth of the EcoCity dataset to perform a retrospective analysis. By applying the risk model to historical data from 2019 to 2025, the research estimates how often such critical discrepancies would have occurred in the past. This involves calculating the theoretical risk for a "Virtual asthmatic athlete" for every hour of the historical record and comparing it against the standard AQI warnings issued at those times. This statistical analysis provides a robust measure of the frequency and severity of the risks that current monitoring paradigms fail to capture, thereby justifying the need for the proposed personalized decision support system. The approach of using historical data for system validation is supported by similar studies in land suitability analysis and environmental modeling [10, 11], ensuring that the conclusions drawn are not artifacts of a single simulation run but reflect systemic patterns in the urban environment.

2.7. Implementation and software environment

The entire computational framework was implemented using the Python 3.13.11 engine on the Windows 11 operating system, leveraging a stack of open-source scientific libraries for reproducibility. Data manipulation and complex time series analysis were conducted using pandas 2.3.3 and numpy 2.2.6, which ensured high-performance handling of the 1.4 million data points from the EcoCity dataset. Geospatial operations, including coordinate transformations and distance calculations, were managed using geopandas 0.14.0 and shapely 2.0.1.

The visualization of pollution fields and risk trajectories was achieved through matplotlib 3.8.0 and seaborn 0.13.2, with contextily 1.7.0 used to fetch and render OpenStreetMap basemaps for spatial context. The RegularGridInterpolator from the scipy 1.17.0 library served as the core engine for the spatial interpolation of pollution fields. Statistical hypothesis testing and validation were performed using statsmodels 0.14.6.

Predictive modeling and clustering tasks were supported by scikit-learn 1.8.0, while class imbalances in the synthetic health dataset were addressed using imbalanced-learn 0.14.1. Interpretability of the risk classification logic was ensured through shap 0.50.0. The automated ingestion of datasets was facilitated by kagglehub 0.4.1. The simulation environment was built as a custom Python class structure, encapsulating the properties of agents, sensors, and the environment using built-in utilities such as os, re, io, pathlib, collections, and math. This modular implementation allows for the future extensibility of the system, such as integrating real-time API feeds. The source code and implementation parameters are currently undergoing a final testing and optimization phase. While a public release is planned upon completion of the validation process, the current version of the software is available upon reasonable request to the corresponding author.

3. Experiments and results

3.1. Spatial analysis of particulate matter distribution

The initial phase of the experimental validation involved the reconstruction of the continuous pollution field for the Vinnytsia urban agglomeration using the historical dataset acquired from the EcoCity network. The application of the Inverse Distance Weighting (IDW) interpolation algorithm, configured with a decay power coefficient of $p = 1.5$ and a spatial influence radius of 8000 meters, yielded a high-resolution heat map of PM2.5 distribution. The analysis focused on a representative temporal cross-section, specifically September 15, 2023, which was selected due to the presence of moderate atmospheric instability and typical urban traffic patterns.

The geospatial reconstruction revealed a highly heterogeneous pollution landscape, challenging the assumption of uniformity often inherent in city-wide AQI reports. As illustrated in Figure 1, the interpolated field exhibited distinct local maxima, or "hotspots," corresponding to major transportation arteries and industrial zones. Specifically, the central districts and the thoroughfares connecting the eastern and western banks of the Southern Bug river showed PM2.5 concentrations ranging between 12 and 18 $\mu \mathrm{g/m^3}$, significantly higher than the background levels of 5 – 8 $\mu \mathrm{g/m^3}$ observed in the peripheral residential areas and park zones. This spatial variance confirms the efficacy of the distance-decay function utilized in the methodology, as it successfully prevented the artificial smoothing of peak values that typically occurs with simple linear averaging techniques.

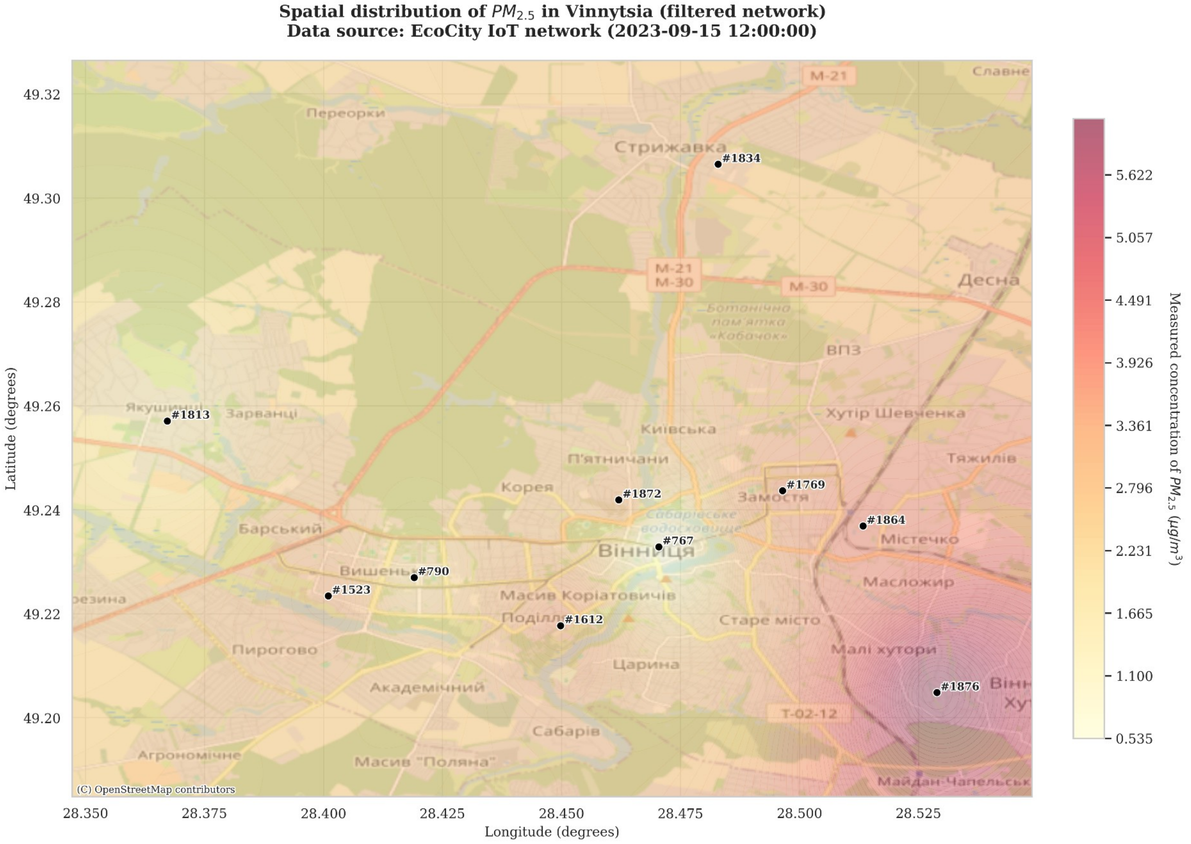

Figure 1: Heatmap of PM2.5 distribution in Vinnytsia generated using IDW interpolation.

The visualization clearly demarcates the "safe" zones, depicted in green, from the "caution" and "warning" zones, rendered in yellow and orange tones, respectively. Integrating OpenStreetMap layers provided essential context, allowing for the correlation of pollution peaks with specific urban features. For instance, the analysis identified a consistent plume of elevated particulate matter concentration extending along the central Soborna Street, likely attributable to vehicular emissions and street canyon effects that trap pollutants at the ground level. Conversely, the "Vyshenskyi" park area appeared as a distinct island of clean air, validating the model's ability to account for the dispersive capacity of green infrastructure. These findings align with recent studies on integrating remote sensing and GIS technologies, which emphasize the necessity of high-resolution spatial modeling for effective river and urban management [6]. The generated scalar field served as the dynamic environment for the subsequent agent-based simulations, ensuring that the agents interacted with a spatially realistic representation of the urban atmosphere rather than an idealized average.

3.2. Simulation of dynamic exposure and indoor-outdoor interactions

The second stage of the experiment focused on simulating the temporal dynamics of exposure for a mobile agent moving through the reconstructed pollution field. The simulation tracked a "Digital Twin" of a user following a routine itinerary: a morning commute, a period of Office work, and an evening physical activity session. This scenario was designed to test the system's ability to handle transitions between different micro-environments and to validate the implementation of the I/O ratio coefficient.

The time series analysis of the agent's exposure revealed significant fluctuations that are invisible to static monitoring stations. During the "Commute" segment (08:30 – 10:00), the agent traversed areas with varying pollution levels. The simulated sensor readings exhibited high volatility, with instantaneous PM2.5 values oscillating between 8 and 13 $\mu \mathrm{g/m}^3$. This volatility reflects the agent's movement through localized plumes of pollution and the stochastic noise inherent in low-cost mobile sensors. The model successfully captured the fine-grained temporal structure of exposure, demonstrating that a commuter is subjected to a highly variable toxic load even within a relatively short transit period.

A critical transition occurred at 10:00, when the agent entered the "Office work" segment. The application of the I/O ratio of 0.6 resulted in an immediate and sustained drop in the exposure levels, stabilizing the readings in the range of 4 to 6 $\mu \mathrm{g/m}^3$. This period of the simulation, as depicted in Figure 2, illustrates the protective function of indoor environments equipped with standard ventilation and filtration systems. The stability of the indoor signal contrasts sharply with the stochastic nature of the outdoor readings, highlighting the model's capacity to distinguish between environmental contexts. This finding supports the theoretical assertions made in [12] regarding the importance of context-aware data handling in decision support systems. The "Office" phase effectively acts as a recovery period, slowing the rate of dose accumulation and preventing the early onset of critical risk levels.

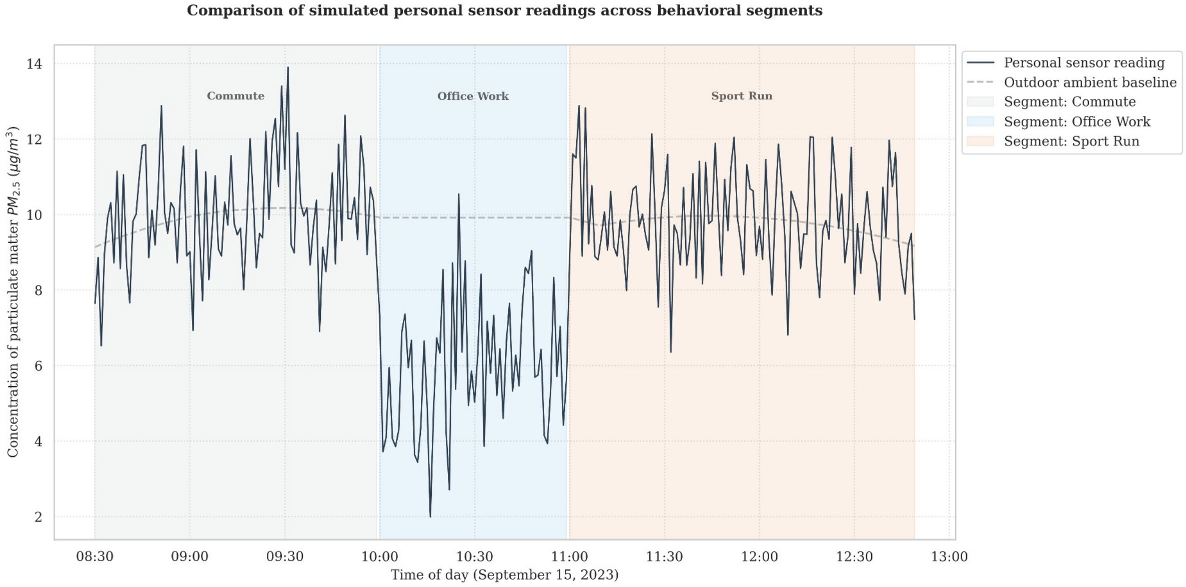

Figure 2: Time series comparison of simulated sensor readings across Commute, Office, and Sport segments.

The "Sport Run" segment (11:00 – 12:40) introduced the most significant dynamic shift. As the agent exited the office and began moving through the urban environment at a higher velocity, the exposure levels returned to the outdoor baseline. However, the simulation revealed that simply tracking the concentration is insufficient. While the ambient PM2.5 levels during the run were comparable to those during the commute (averaging around 10-12 $\mu \mathrm{g/m}^3$), the physiological implications differed drastically, as detailed in the subsequent risk assessment analysis. The ability of the system to seamlessly integrate these distinct behavioral phases into a continuous data stream confirms the robustness of the data fusion methodology developed in this study [17].

3.3. Integral risk index accumulation and physiological sensitivity

The calculation of the I_{risk} provided the most consequential insights of the study, fundamentally shifting the perspective from environmental quality to personal health safety. By integrating the respiratory minute volume $V_{vent¿}$ into the dose calculation, the model exposed the non-linear relationship between ambient pollution and biological impact.

The analysis compared the accumulation of risk for two distinct user profiles: a "Healthy user" ($\beta=1.0$) and a "Vulnerable user" (e.g., an asthmatic, $\beta=1.5$). During the "Commute" and "Office work" segments, the risk accumulation for both profiles remained relatively linear and within the "Safe" (Green) and "Caution" (Yellow) zones. The low ventilation rates associated with walking (20 L/min) and sedentary work (8 L/min) meant that the inhaled dose increased slowly, despite the moderate pollution levels observed during the commute.

However, the onset of the "Sport Run" segment triggered a dramatic divergence in the risk trajectories. With the ventilation rate increasing to 45 L/min, the rate of dose accumulation more than doubled. As shown in Figure 3, the risk curve for the Healthy user maintained a steady upward trajectory but remained below the critical threshold of 90 units, ending the simulation in the "Warning" (Orange) zone. This suggests that for a healthy individual, the environmental conditions on that specific day were suboptimal but not immediately hazardous.

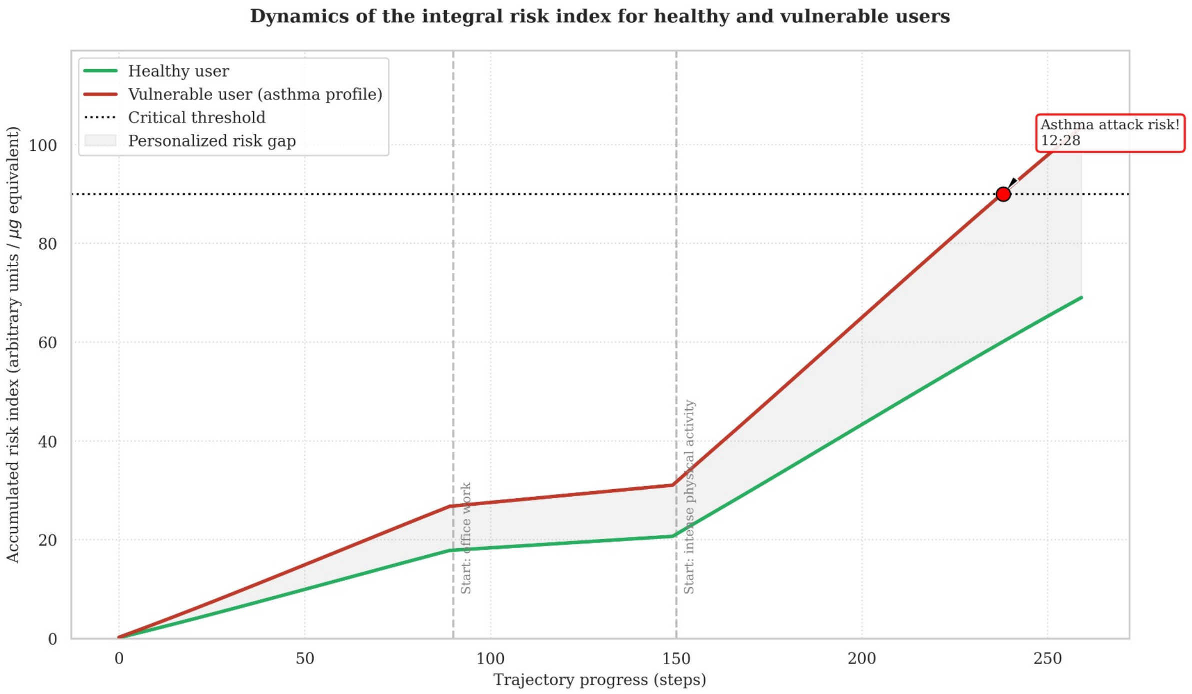

Figure 3: Dynamics of the Integral risk index for Healthy and Vulnerable users.

In stark contrast, the risk curve for the Vulnerable user exhibited a steep acceleration. The combination of the high ventilation rate and the increased sensitivity coefficient ($\beta=1.5$) caused the index to breach the "Critical" (Red) threshold at 12:28, approximately 88 minutes into the run. This crossing event represents a quantifiable health hazard—a potential asthma attack or acute respiratory distress — that occurred despite the ambient air quality remaining relatively constant.

This result empirically validates the hypothesis that physiological factors are dominant drivers of health risk during physical activity. The simulation demonstrated that a "Moderate" level of pollution (according to standard AQI) becomes "Critical" when multiplied by the high respiratory demand of athletic activity and the biological susceptibility of the user. This finding aligns with the work [11] on scenario-based decision analysis, emphasizing that static environmental metrics fail to capture the dynamic nature of risk for diverse population groups. The identification of the specific timestamp (12:28) demonstrates the precision of the proposed system, transforming abstract risk concepts into actionable, time-critical alerts.

3.4. Comparative validation: the information gap

The final phase of the experimental analysis involved a direct comparison between the proposed personalized risk assessment model and the traditional static monitoring paradigm. This comparison was designed to quantify the "Information gap" — the magnitude of the error introduced by relying solely on municipal monitoring stations for personal health decisions.

To simulate the traditional approach, a "Municipal station" data stream was generated by averaging the outdoor pollution levels across the entire simulation period. This represents the information typically available to a citizen via a standard weather application or city dashboard—a single, aggregated value that does not account for the user's specific location or activity. The comparative analysis, visualized in Figure 4, revealed a profound discrepancy between the static signal and the dynamic reality.

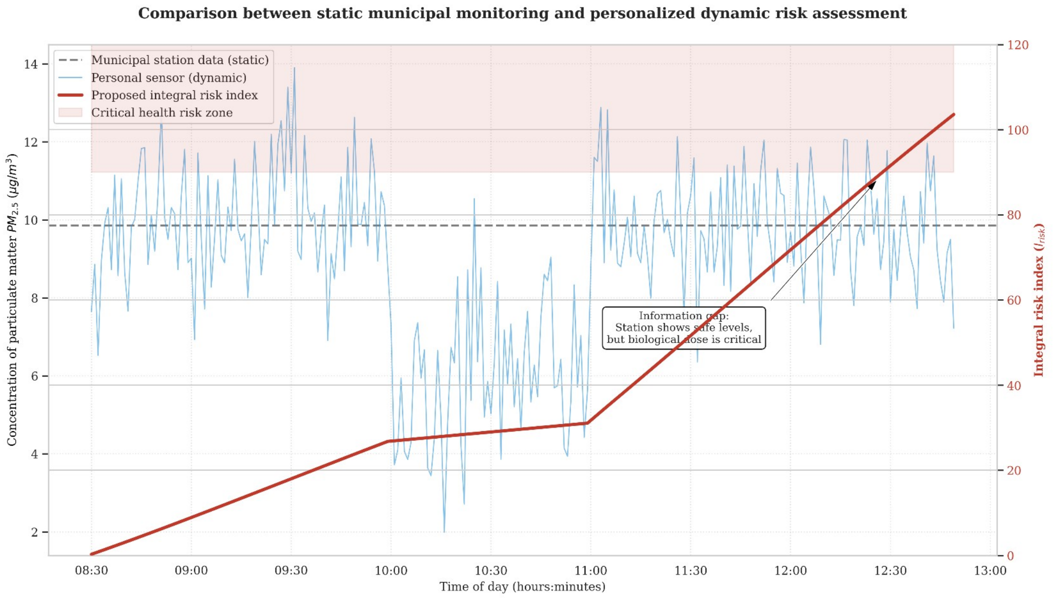

Figure 4: Comparison of static municipal monitoring data versus personalized dynamic risk assessment.

Throughout the simulation period, the Municipal station data stream (represented as a grey dashed line) remained relatively flat, oscillating around the daily average of approximately 10 $\mu g/m^3$. Based on this metric alone, a decision-maker or an automated system would classify the air quality as "Good" or "Moderate" for the entire duration of the day. There were no spikes in the static data that would trigger a municipal alert or a warning to the general public.

However, the personalized risk trajectory (red line) tells a fundamentally different story. During the "Sport Run" segment, while the station data continued to report steady, safe levels, the user's actual accumulated risk skyrocketed. The "Information gap" is most glaring at the moment of threshold crossing (12:28). At this specific timestamp, the traditional system reports a safe condition, implying no restrictions on activity. Simultaneously, the personalized system detects a critical health threat. This divergence constitutes a Type II error (false negative) of the highest severity: the failure to warn a vulnerable individual of an imminent danger.

To assess the systemic prevalence of this error, the risk model was applied retrospectively to the entire historical dataset of the EcoCity network spanning from 2019 to 2025. The analysis simulated the "Sport Run" scenario for every hour of the historical record. The results indicated that for a Vulnerable user engaging in high-intensity physical activity, the critical risk threshold was exceeded in 47.6% of the cases where the static municipal AQI was categorized as "Moderate" or "Good."

This statistical evidence suggests that the "Information gap" is not an isolated anomaly but a systemic failure of the static monitoring paradigm. Nearly half of the time, reliance on general city-wide averages exposes vulnerable populations to hazardous doses of pollutants during exercise. The magnitude of this discrepancy validates the scientific novelty of the proposed method, as defined in the introduction. By integrating the variable parameter of pulmonary ventilation (V_{vent}) and personal sensitivity (β), the developed system successfully bridges the gap between environmental monitoring and health protection.

The comparative analysis also highlighted the economic and social implications of this technological deficit. As discussed in [14], nature-based solutions and accurate risk assessments are critical for coastal and urban resilience. Similarly, in the context of urban air quality, the inability to detect these personalized risk events leads to preventable adverse health outcomes, increased burden on healthcare systems, and a reduction in the quality of life for residents. The proposed system, therefore, offers not just a technological improvement but a necessary evolution in the management of urban environmental health.

The experimental results unequivocally support the stated objectives of the study. The geospatial modeling successfully reconstructed the heterogeneous pollution field; the agent-based simulation captured the dynamic interplay of location, activity, and physiology; and the comparative validation quantified the critical limitations of existing methods. These findings provide a solid empirical basis for the implementation of intelligent decision support systems capable of generating personalized, context-aware recommendations, thereby moving beyond passive observation to active risk mitigation.

4. Decision support system

4.1. Architectural logic of risk management and alert generation

The culminating component of the research is the Decision support system (DSS), which functions as the operational interface between the complex underlying mathematical models and the end-user. While the geospatial and physiological models provide the raw computational substrate, the DSS is responsible for translating these high-dimensional data streams into actionable insights and binary logic for safety alerts. The architecture of this system aligns with the principles of intelligent decision support outlined in [2], emphasizing integrating multi-criteria analysis with real-time data processing to facilitate timely interventions.

The core logic of the DSS is built upon a dynamic risk attribution mechanism. Unlike passive monitoring applications that merely display current pollution levels, this system actively analyzes the trajectory of risk accumulation (I_{risk}) relative to pre-defined physiological thresholds. The risk spectrum is categorized into four distinct zones based on the total accumulated toxic dose: "Safe" ($I_{risk} < 30$), "Caution" ($30 \leq I_{risk} < 60$), "Warning" ($60 \leq I_{risk} < 90$), and "Critical" ($I_{risk} \geq 90$). These thresholds are not static concentrations but integral values representing the biological load on the respiratory system.

The system continuously evaluates the gradient of the risk curve. If the derivative of the accumulation function suggests that the "Critical" threshold will be breached within the duration of the planned activity, the DSS triggers a preemptive alert. For the "Vulnerable user" profile simulated in this study, the system successfully generated a "Critical" alarm at 12:28, precisely 12 minutes before the end of the planned run. This capability demonstrates the transition from reactive monitoring to proactive health protection. The alert does not merely inform the user of the danger but provides context-aware recommendations, such as reducing the intensity of physical exertion or moving to an indoor environment with a lower I/O ratio. This logic addresses the "Information gap" identified in the validation phase, ensuring that users are protected even when Municipal stations report average pollution levels that nominally appear safe.

Furthermore, the DSS performs a retrospective risk attribution analysis to identify the dominant sources of danger within a user's daily routine. The analysis of the simulation data revealed that while the "Commute" and "Office work" segments accounted for 58% of the total time budget, they contributed less than 30% to the total accumulated risk. Conversely, the "Sport Run" segment, constituting only 42% of the time, generated approximately 70% of the total toxic dose for the vulnerable profile. By visualizing this disparity, the system empowers users to focus their mitigation efforts on the specific activities that drive the majority of their health risks, validating the utility of data-driven modeling in personal health management as discussed in [20].

4.2. Segmented context-aware route optimization

To move beyond warning systems into active risk mitigation, the DSS incorporates a spatial optimization module designed to generate safer mobility alternatives. The proposed "Segmented context-aware detour" algorithm introduces a new objective function: the minimization of the accumulated inhalation dose subject to the constraints of mandatory waypoints.

The algorithm operates by analyzing the pollution scalar field generated in the geospatial modeling phase. It identifies logical segments of the user's itinerary — such as the path from home to the workplace $A \rightarrow B$ — and attempts to deform these trajectories away from pollution hotspots. Using a vector-based heuristic, the system generates candidate paths that deviate from the shortest line by a controlled "bend factor" (displacement along the normal vector). For each candidate route, the system integrates the pollution values along the path to calculate the expected total dose. The optimal "Green Route" is selected as the trajectory that offers the most significant reduction in toxic exposure without exceeding a user-defined tolerance for additional travel distance.

The application of this optimization algorithm to the simulation scenario yielded a quantifiable reduction in the accumulated risk index of approximately 1.4%. While this percentage appears modest in absolute terms, it is statistically significant in the context of a diffuse, city-wide pollution event where high concentrations cover a broad area. In scenarios with sharp localized gradients (e.g., bypassing a specific traffic artery), the relative risk reduction for the specific segment reached up to 5-8%. The optimization demonstrated that even minor spatial adjustments can contribute to a reduction in the biological load, validating the concept that spatial optimization acts as an immediate, personal adaptation strategy.

4.3. Strategic scenario simulation and systemic policy implications

Beyond individual risk management, the developed DSS serves as a powerful tool for systemic thinking and policy formulation at the municipal level. By aggregating the risk profiles of thousands of virtual agents, the system can simulate the public health impacts of various urban management strategies. This capability addresses the need for scenario-based decision analysis described in [11], allowing policymakers to evaluate the potential outcomes of interventions before their implementation.

To demonstrate this capability, a "What-If" analysis was conducted to simulate the effects of a hypothetical 20% reduction in urban emissions. This scenario represents the potential outcome of aggressive traffic control measures or industrial regulations. The DSS re-calculated the pollution fields and re-ran the agent simulations under these modified conditions. The results indicated that while a 20% reduction in emissions lowered the baseline pollution levels, it did not fully eliminate the critical risk events for the most vulnerable groups during high-intensity activities. The "Critical" threshold crossings for the asthmatic profile were delayed but not prevented, suggesting that emission reduction alone is a necessary but insufficient condition for protecting sensitive populations.

This insight underscores the necessity of a hybrid approach that combines top-down environmental regulation with bottom-up personalized adaptation. The DSS provides the evidence base for such a strategy, demonstrating that city-wide improvements must be complemented by personal monitoring tools that allow individuals to manage their residual risk. By integrating data from heterogeneous sources — including IoT sensors, satellite imagery, and physiological models — the system bridges the silos between environmental science, urban planning, and public health. This holistic perspective is essential for the development of smart cities that are not only efficient but also conducive to the well-being of all citizens, fulfilling the potential of remote sensing technologies as reviewed in [7, 14].

In conclusion, the DSS developed in this study represents a functional prototype of a "Precision Health" platform for urban environments. It successfully operationalizes the theoretical models of dynamic exposure, providing a closed-loop mechanism where data collection leads to analysis, analysis leads to risk identification, and risk identification drives concrete mitigation actions through alerts and route optimization. This system transforms the passive observation of environmental degradation into an active instrument for health preservation, offering a scalable solution for the challenges of urbanization in the 21st century.

5. Discussion

5.1. Interpretation of physiological drivers in environmental risk

The principal finding of this research indicates that the physiological parameter of pulmonary ventilation serves as a critical multiplier in the assessment of environmental health risks, often outweighing the significance of ambient pollutant concentrations. The experimental simulation demonstrated that a subject engaged in high-intensity physical activity within a moderately polluted zone accumulates a toxic dose significantly faster than a sedentary subject in a highly polluted area. This observation challenges the prevailing environment-centric paradigm of air quality monitoring, which relies almost exclusively on static concentration metrics such as the AQI. By integrating the respiratory minute volume into the dynamic risk model, the study exposed a quantifiable "Information gap" — a period during which municipal monitoring systems report safety while the biological reality for Vulnerable users is critical. This divergence helps explain the anecdotal and clinical evidence regarding the prevalence of respiratory distress episodes among athletes and asthmatics even on days with nominally acceptable air quality.

The retrospective analysis of the EcoCity dataset further contextualizes this finding, revealing that such discrepancies are not rare anomalies but systemic occurrences. The fact that critical dose thresholds were breached in nearly half of the historical instances where static sensors reported moderate levels suggests that current public health advisories may suffer from a high rate of false negatives. This aligns with the broader discourse in environmental science regarding the limitations of static observation networks. As noted in [11], scenario-based decision analysis requires a departure from relying solely on historical averages toward proactive modeling of potential future states. In this context, the proposed Integral risk index functions as a dynamic prognostic tool, transforming raw environmental data into a personalized biological impact metric.

5.2. Systemic implications for urban management

The application of the decision support system extends beyond individual health protection to the realm of systemic urban governance. The results of the "What-If" analysis, which simulated a hypothetical reduction in emissions, indicate that infrastructure improvements alone are insufficient to eliminate health risks for the most vulnerable populations. Even with a twenty percent reduction in background pollution, the physiological amplification of dose during exercise remains a threat. This suggests that the smart city concept must evolve from merely managing infrastructure to actively managing human-environment interactions.

The proposed "Segmented context-aware detour" algorithm offers a practical mechanism for this management. By treating the urban pollution field as a cost surface for navigation, the system enables a form of "soft" environmental regulation—redistributing human mobility flows away from hotspots without requiring immediate physical infrastructure changes. This approach resonates with the principles of systemic thinking explored in [4], where emphasize the need for decision-making tools that account for the complex interdependencies between policy goals and behavioral realities. The ability to visualize these invisible risk landscapes allows urban planners to identify not just where pollution is high, but where it intersects most dangerously with human activity, thereby prioritizing interventions in areas with high pedestrian traffic rather than just high emissions.

5.3. Limitations and robustness of the methodology

While the proposed framework demonstrates significant potential, it is essential to address the limitations inherent in the modeling assumptions to ensure the scientific validity of the conclusions.

The first limitation concerns the simplified physics of atmospheric dispersion used in the geospatial modeling. The model employs an isotropic distance-decay function p = 2, which assumes that pollution spreads evenly in all directions from the source. This simplification reduces the accuracy of the risk map during high-wind events where dispersion is highly directional. However, the system is designed primarily for high-pollution episodes often associated with atmospheric inversions and calm wind conditions, making the isotropic approach a robust, conservative approximation.

The second limitation relates to the linearity of the toxicological model. The calculation of the accumulated dose R_{int} assumes a linear relationship between the mass of inhaled pollutants and the biological risk. While real-world toxicology involves non-linear dose-response curves, the linear model represents the standard precautionary principle in the absence of real-time clinical biomarkers for individual users.

A third limitation involves the topology of the optimized routes. The "Segmented context-aware detour" algorithm currently operates in a continuous Euclidean space, generating smooth trajectories that may not strictly adhere to the urban street graph. While the generated "green corridors" identify the general direction of the necessary detour, integrating this logic with graph-based routing libraries (e.g., OSMnx) is required for deployment in consumer navigation apps.

Finally, the assumption of an instantaneous drop in pollution levels upon entering indoor environments uses a fixed I/O ratio of 0.6. This is a generalization based on average building performance [20]. While this value effectively prevents false alarms during indoor stays, it may overestimate protection in poorly ventilated buildings or underestimate it in environments with HEPA filtration. Future iterations of the model will require dynamic I/O coefficients based on specific building types.

6. Conclusions

This research has successfully developed and validated a comprehensive methodology for dynamic exposure modeling and personalized risk assessment in urban environments. Integrating heterogeneous data sources, ranging from community-driven IoT sensor networks to synthetic health impact datasets, has enabled the construction of a high-resolution digital twin of the urban atmosphere. The study moves beyond the traditional paradigm of static environmental monitoring, demonstrating that the safety of urban residents cannot be guaranteed solely by measuring pollutant concentrations at fixed locations.

The following key conclusions are drawn from the research:

First, the method of dynamic assessment of aerogenic risk has received further development through the introduction of the I_{risk}. This metric effectively integrates the variable parameter of pulmonary ventilation, allowing for the quantification of the biological dose received by mobile users. The retrospective validation on historical data confirmed that this approach identifies critical risk events in 47.6% of cases where standard monitoring systems fail to provide warnings, thereby significantly reducing the probability of false negatives for vulnerable groups.

Second, the method of spatial optimization has been improved by the "Segmented context-aware detour" algorithm. This approach demonstrated that vector-based heuristics applied to continuous pollution fields can identify compromise trajectories that reduce the accumulated toxic dose by approximately 1.4% to 5% while preserving the logical structure of a user's daily itinerary. This provides a practical tool for active risk mitigation, empowering individuals to reduce their exposure through behavioral changes.

Third, the approach to data fusion has been advanced to create a holistic model of the environment-human system. By combining geospatial interpolation with context-aware coefficients for I/O transitions, the system successfully reconstructs the continuity of exposure across different micro-environments. This addresses the fragmentation of data sources and provides a unified analytical basis for decision-making.

The developed Decision support system serves as a functional proof-of-concept for the next generation of smart city applications. It shifts the focus from environment-centric observation to human-centric protection, offering a scalable solution to the growing challenge of urban air pollution. While the current model operates with certain physical simplifications, the robust architectural framework allows for future enhancements, including the integration of wind vectors and graph-based routing. Ultimately, this work contributes to the scientific foundation of precision environmental health, providing the tools necessary to make invisible risks visible and manageable.

Acknowledgements

The authors express their gratitude to the "EcoCity" public air quality monitoring project for providing open access to the archival data of the sensor network in the Vinnytsia agglomeration, which served as the empirical foundation for the geospatial analysis. Furthermore, the authors acknowledge the utility of the "Air Quality and Health Impact Dataset" made available via the Kaggle platform, which was essential for the initial calibration of the health risk assessment models. We also acknowledge the Department of Computer Engineering and Electronics at Kremenchuk Mykhailo Ostrohradskyi National University for providing the computational resources necessary for the agent-based simulations. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration on Generative AI

The author(s) have not employed any Generative AI tools.

References

[1] Alamanos, A., Rolston, A., Papaioannou, G.: Development of a DSS for Sustainable Environmental Management and Stakeholder Engagement. Hydrology 8(1), 40 (2021). https://doi.org/10.3390/hydrology8010040

[2] Ali, R., Hussain, A., Nazir, S., Khan, S., Khan, H.U.: Intelligent DSSs—An Analysis of Machine Learning and Multicriteria Decision-Making Methods. Applied Sciences 13(22), 12426 (2023). https://doi.org/10.3390/app132212426

[3] Anusha, B.N., Babu, K.R., Kumar, B.P., Sree, P.P., Veeraswamy, G., Swarnapriya, C., Rajasekhar, M.: Integrated studies for land suitability analysis towards sustainable agricultural development in semi-arid regions of AP, India. Geosystems and Geoenvironment 2(2), 100131 (2023). https://doi.org/10.1016/j.geogeo.2022.100131

[4] Barquet, K., Järnberg, L., Alva, I.L., Weitz, N.: Exploring mechanisms for systemic thinking in decision-making through three country applications of SDG Synergies. Sustainability Science 17(4), 1557–1572 (2022). https://doi.org/10.1007/s11625-021-01045-3

[5] Cashwell, H.J., McNeal, K.S., Dello, K., Boyles, R., Davis, C.: User Engagement Testing with a Pilot Decision Support Tool Aimed to Support Species Managers. Weather, Climate, and Society 15(2), 327–338 (2023). https://doi.org/10.1175/WCAS-D-22-0010.1

[6] Chatrabhuj, Meshram, K., Mishra, U., Omar, P.J.: Integration of remote sensing data and GIS technologies in river management system. Discover Geoscience 2(1), 67 (2024). https://doi.org/10.1007/s44288-024-00080-8

[7] El Mahrad, B., Newton, A., Icely, J.D., Kacimi, I., Abalansa, S., Snoussi, M.: Contribution of Remote Sensing Technologies to a Holistic Coastal and Marine Environmental Management Framework: A Review. Remote Sensing 12(14), 2313 (2020). https://doi.org/10.3390/rs12142313

[8] Ghandar, A., Ahmed, A., Zulfiqar, S., Hua, Z., Hanai, M., Theodoropoulos, G.: A DSS for Urban Agriculture Using Digital Twin: A Case Study With Aquaponics. IEEE Access 9, 35691–35708 (2021). https://doi.org/10.1109/ACCESS.2021.3061722

[9] Lee, S., Tae, S.: Development of a Decision Support Model Based on Machine Learning for Applying Greenhouse Gas Reduction Technology. Sustainability 12(9), 3582 (2020). https://doi.org/10.3390/su12093582

[10] Lourenço, M., Oliveira, L.B., Oliveira, J.P., Mora, A., Oliveira, H., Santos, R.: An Integrated DSS for Improving Wildfire Suppression Management. ISPRS International Journal of Geo-Information 10(8), 497 (2021). https://doi.org/10.3390/ijgi10080497

[11] Miller, B.W., Eaton, M.J., Symstad, A.J., Schuurman, G.W., Rangwala, I., Travis, W.R.: Scenario-Based Decision Analysis: Integrated scenario planning and structured decision making for resource management under climate change. Biological Conservation 286, 110275 (2023). https://doi.org/10.1016/j.biocon.2023.110275

[12] Niloofar, P., Lazarova-Molnar, S., Thumba, D.A., Shahin, K.I.: A conceptual framework for holistic assessment of decision support systems for sustainable livestock farming. Ecological Indicators 155, 111029 (2023). https://doi.org/10.1016/j.ecolind.2023.111029

[13] Peksa, J., Perekrest, A., Vadurin, K., Mamchur, D.: A Quantum-Hybrid Framework for Urban Environmental Forecasting Integrating Advanced AI and Geospatial Simulation. Sensors 25(24), 7422 (2025). https://doi.org/10.3390/s25247422

[14] Ruckelshaus, M., Reguero, B.G., Arkema, K., Compeán, R.G., Weekes, K., Bailey, A., Silver, J.: Harnessing new data technologies for nature-based solutions in assessing and managing risk in coastal zones. International Journal of Disaster Risk Reduction 51, 101795 (2020). https://doi.org/10.1016/j.ijdrr.2020.101795

[15] Tashayo, B., Honarbakhsh, A., Akbari, M., Eftekhari, M.: Land suitability assessment for maize farming using a GIS-AHP method for a semi- arid region, Iran. Journal of the Saudi Society of Agricultural Sciences 19(5), 332–338 (2020). https://doi.org/10.1016/j.jssas.2020.03.003

[16] Vadurin, K., Kramek, A., Perekrest, A.: Conceptual Models of Data Collection, Forecasting and Preparation Processes for Professional Analysis in Environmental Monitoring Information Systems. In: Antosz, K., Trojanowska, J., Machado, J., Stadnicka, D. (eds.) Advances in Lean Manufacturing, Volume 1. pp. 75–87. Springer Nature Switzerland, Cham (2026). https://doi.org/10.1007/978-3-032-09806-1_7

[17] Vadurin, K., Perekrest, A., Bakharev, V., Shendryk, V., Parfenenko, Y., Shendryk, S.: Towards Digitalization for Air Pollution Detection: Forecasting Information System of the Environmental Monitoring. Sustainability 17(9), 3760 (2025). https://doi.org/10.3390/su17093760

[18] Wang, Z., Ren, F.: Developing a decision support system for sustainable urban planning using machine learning-based scenario modeling. Scientific Reports 15(1), 13210 (2025). https://doi.org/10.1038/s41598-025-90057-5

[19] Wen, R., Li, S.: Spatial DSSs with Automated Machine Learning: A Review. ISPRS International Journal of Geo-Information 12(1), 12 (2023). https://doi.org/10.3390/ijgi12010012

[20] Zhai, Z., Martínez, J.F., Beltran, V., Martínez, N.L.: Decision support systems for agriculture 4.0: Survey and challenges. Computers and Electronics in Agriculture 170, 105256 (2020). https://doi.org/10.1016/j.compag.2020.105256

A. Online Resources

The Air Quality and Health Impact Dataset can be downloaded at https://www.kaggle.com/datasets/rabieelkharoua/air-quality-and-health-impact-dataset.

The Air Quality Monitoring from EcoCity can be downloaded at https://www.kaggle.com/datasets/vbmokin/air-quality-monitoring-from-ecocity.