

Health Aware Optimized Route Planner for Reduced Urban Hyperlocal Pollution Exposure

Reem Abdelaziz Alshamsi¹, Luca Foschini², Isam Mashhour Al Jawarneh¹

¹ Department of Computer Science, University of Sharjah, Sharjah P. O. Box 27272, United Arab Emirates

² Dipartimento di Informatica – Scienza e Ingegneria, University of Bologna, Viale Risorgimento 2, 40136 Bologna, Italy
U23102393@sharjah.ac.ae, luca.foschini@unibo.it, ijawarneh@sharjah.ac.ae

Abstract—This paper presents a routing engine designed to direct users through routes with healthier air quality in urban environments. Employing an approach that integrates geospatial analysis and air quality indices, we establish 'healthy routes' that minimize exposure to harmful particulate matter (PM). Our methodology leverages shortest path algorithms, enhanced by air quality ratings, to generate environmentally optimized paths. Initial results indicate identical routes using PM2.5 and NO2 data, suggesting effective route consistency. This study aims to contribute to urban planning and public health by providing a practical tool for minimizing pollution exposure.

Keywords— *Urban Air Quality; Routing Engine; Particulate Matter; Geospatial Analysis; Pollution Exposure*

I. INTRODUCTION

Particulate matter (PM) pollution is a significant environmental and health concern, as highlighted by its adverse impacts on respiratory health, especially in children, with biomass fuel emissions contributing to a notable increase in bacterial pneumonia deaths [1]. PM exposure triggers severe inflammatory responses in the lungs, exacerbated by various PM types such as silica and cigarette smoke, leading to oxidative stress and activation of critical inflammatory pathways [2]. Moreover, PM is implicated in cancer development and progression through epigenetic mechanisms, with both PM10 and PM2.5 inducing systemic inflammation and oxidative stress, elevating risks of cardiovascular diseases and cancer [3]. Urban PM primarily originates from industrial emissions, vehicle exhaust, and natural sources like dust, with secondary aerosols forming through reactions between primary pollutants and gases like SO2 and NOx, further deteriorating air quality. Regulatory standards emphasize the necessity of controlling PM sizes and reducing levels to mitigate associated health risks. Deposition of PM within the respiratory system varies by particle size, influencing diverse immunologic and inflammatory responses and impacting critical respiratory functions. The elevated toxicity risk posed by ultrafine particles due to their size and surface area

underscores the diverse chemical and physical composition of PM and its profound health ramifications.

In response to the pressing concern of particulate matter (PM) pollution in the air, this paper introduces a solution: a routing engine designed to navigate users along healthier paths based on PM ratings ranging from 1 to 5, where 1 denotes healthy air quality and 5 signifies elevated PM levels posing health risks. The study focuses on analyzing New York City (NYC) air quality data, where 18 alternative "healthy" routes were meticulously crafted considering both PM2.5 and NO2 data. Initially, the shortest path algorithm was utilized to establish routes based on distance, followed by the generation of new paths using either the average PM2.5 or NO2 ratings for specific geohashes.

This paper comprises the following sections: I - Introduction, II - Preliminaries (covering PM and NO2 pollution sources, health impacts, and regulatory standards), III - GreenRouter: Health-Aware Route Planner (methodology: detailing data collection, preprocessing, and route generation), IV - Experimental Evaluation, V - Literature (examining urban planning integration of environmental data and Spatial Data Processing Techniques), and VI - Conclusion & Future Works.

II. PRELIMINARIES

In this section, we briefly discuss the preliminaries and theory background of relevant topics that are necessary to comprehend subsequent discussion about the design and implementation of our novel system in the paper.

A. Particulate Matter (PM)

Particulate matter (PM) is a critical air pollutant, classified into PM10 and PM2.5, with PM2.5 being particularly harmful due to its ability to penetrate deeply into the respiratory system and cause serious health issues, including cardiovascular and respiratory diseases, and lung cancer. The most damaging effects are attributed to the primary combustion-derived nanoparticulate component of PM2.5 [4]. Urban PM

originates from industrial emissions, vehicle exhaust, and natural sources like dust. Secondary aerosols form through reactions between primary pollutants and gases such as SO₂ and NO_x, exacerbating air quality degradation. Regulatory standards emphasize the importance of controlling PM sizes and reducing levels to mitigate health risks [5]. The deposition of PM within the respiratory system varies by particle size, affecting different areas from the trachea to the alveoli, leading to diverse immunologic and inflammatory responses, impacting alveolar macrophage function, and other critical respiratory functions. PM underscores the elevated toxicity risk posed by ultrafine particles, attributing it to their size and surface area. This underscores PM's diverse chemical and physical composition and its profound health ramifications.

B. Health Impacts of Particulate Matter (PM)

Exposure to particulate matter (PM) significantly increases the risk of respiratory infections, particularly in children, with notable impacts from biomass fuels leading to a substantial rise in bacterial pneumonia deaths. [1] PM exposure is also associated with severe inflammatory responses in the lungs, where various PM types, such as those from silica and cigarette smoke, elevate cytokine and chemokine production, causing oxidative stress and activating critical inflammatory pathways mediated by Type II epithelial cells. [2]

Furthermore, PM has profound implications for cancer development and progression through its influence on gene regulation via epigenetic mechanisms. Both PM₁₀ and PM_{2.5} are known to induce systemic inflammation and oxidative stress, linked to increased risks of cardiovascular diseases and cancer. The activation of proinflammatory pathways by PM contributes to widespread cellular and tissue damage, highlighting the severe health risks associated with particulate air pollution. These effects collectively underscore the dangers of PM, emphasizing the need for effective air quality management to mitigate its health impacts [3].

III. GREENROUTER: HEALTH-AWARE ROUTE PLANER

In this section, we show the design and implementation of our novel system GreenRouter for the efficient generation of region-based geo-maps from big geotagged locational data, with a specific focus in this paper on choropleth maps.

The methodology involves key steps: data collection, exploratory data analysis (EDA), preprocessing, sampling, experimentation, and evaluation. We adopt

a comprehensive approach integrating these steps to develop and assess pollution-aware routing solutions. Beginning with thorough data collection, we gathered detailed air quality metrics from New York City, including PM_{2.5} and NO₂ levels recorded between September and December 2021. This data offers crucial temporal and spatial insights essential for precise route optimization.

Data collection. Two hyperlocal air quality datasets from New York City, covering September to December 2021, were utilized. The first dataset focuses on PM_{2.5} measurements, while the second centers on NO₂ levels.¹ OpenStreetMap provided graph information to support the implementation of the shortest path algorithm.

Data collection involved gathering comprehensive air quality measurements of PM_{2.5} and NO₂ from various monitoring stations across New York City. Integration with OpenStreetMap enhanced the dataset with geographical context, facilitating precise spatial analysis. This dataset captures pollution level fluctuations across boroughs and time periods, serving as a foundation for realistic urban mobility solutions. High-resolution spatial data from OSM ensured accurate pathfinding by providing detailed urban topography insights.

EDA. During the exploratory data analysis phase, we thoroughly examined the air quality data, assessing the distribution of PM_{2.5} and NO₂ concentrations across New York City's boroughs and throughout the months covered by the dataset. By plotting these distributions, we identified patterns and anomalies in pollution levels, offering insights into environmental impacts at a detailed level. Furthermore, we explored correlations between air quality and meteorological factors like humidity and temperature. This statistical exploration aided our understanding of how environmental conditions influence pollution levels, thereby influencing routing decisions made by our engine.

We have analyzed the distribution of air quality datasets for PM and NO₂ across different boroughs of New York City. Notably, The Bronx exhibits a considerably higher count of both PM and NO₂ data points compared to other boroughs, suggesting a potential focus area for detailed air quality monitoring. The monthly distribution of data records for PM and NO₂ throughout the study period shows that September and October exhibit an equal and peak collection rate of approximately 35%, while records decline notably by December, suggesting a seasonal variation in data collection intensity.

The distributions of temperature and humidity within the PM and NO₂ datasets reveal that the temperature data tends to skew towards moderate values, whereas humidity exhibits a broad distribution,

¹ Dataset link: <https://zenodo.org/records/7961851>

indicating diverse environmental conditions during the data collection period across New York City. For both datasets, the temperature and humidity are the same, as the data was collected around the same period, from the same areas around NYC.

The frequency distribution of PM2.5 concentration values across New York City indicates a skewed pattern, with a prominent peak around the 2-6 $\mu\text{g}/\text{m}^3$ range, suggesting that most recorded PM2.5 concentrations fall within this relatively low and healthier range. Beyond 6 $\mu\text{g}/\text{m}^3$, the distribution gradually tapers off, with fewer data points recorded at higher concentrations.

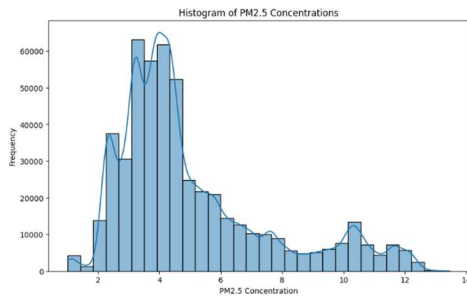


Fig. 1. Histogram of PM 2.5 Concentrations

The histogram shown in Figure 1 highlights the significance of the air quality rating system, where values up to 3 $\mu\text{g}/\text{m}^3$ are assigned a rating of '1' (healthiest), and those exceeding 12 $\mu\text{g}/\text{m}^3$ are assigned a rating of '5' (worst). This categorization serves as a crucial factor in determining routes that minimize exposure to harmful particulate matter. By integrating these ratings into routing algorithms, paths with cleaner air can be prioritized, a strategy particularly beneficial in urban planning and public health contexts.

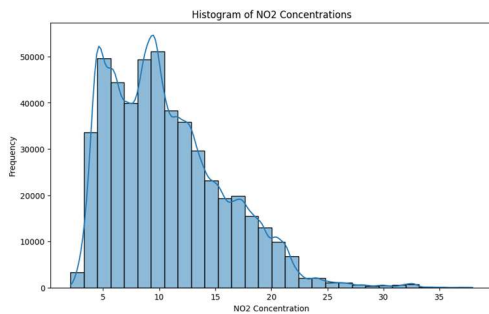


Fig. 2. Histogram of NO2 Concentrations

The histogram depicting NO2 concentration values (Figure 2) reveals a skewed distribution, encompassing a broader range of higher concentration frequencies compared to PM2.5. Peak frequencies are observed between 5 to 15 $\mu\text{g}/\text{m}^3$, indicating a significant portion

of the NO2 data falls within a moderate pollution level. As concentrations exceed 20 $\mu\text{g}/\text{m}^3$, the distribution markedly declines. In line with the applied air quality rating system, NO2 concentrations are categorized on a scale from '1' ($\leq 5 \mu\text{g}/\text{m}^3$, healthiest) to '5' ($> 20 \mu\text{g}/\text{m}^3$, worst). This categorization aids in assessing the associated health risks and facilitates the implementation of a routing engine that dynamically adjusts paths to circumvent areas with high NO2 levels, thus reducing residents' exposure to this pollutant. Such an approach is pivotal for improving urban livability and health outcomes by strategically navigating away from more polluted, potentially hazardous areas.

Preprocessing. Preprocessing the air quality data involved a series of pivotal steps to ready it for analysis. Initially, we converted the raw data into a more manageable format by integrating it with GeoJSON, enriching it with geographical encoding.

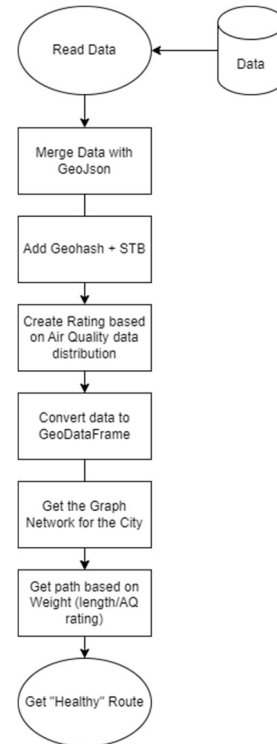


Fig. 3. Process Diagram

This integration facilitated the application of Geohash encoding, enabling precise spatial data partitioning. Each data point was then assigned an air quality rating based on predefined thresholds, pivotal for subsequent routing decisions. This structured and enriched dataset was crucial for the effective application of our routing algorithms, ensuring accurate representation of geographical and environmental factors.

The preprocessing commenced with reading the air quality data in CSV format. Subsequently, the data was merged with GeoJSON to incorporate geographical information from New York City. A Geohash of precision 5 was initially added to the data frame, along with a modified version of the STB² (Space-Time-Box) containing the geohash and time for each record. Next, a rating was generated based on the distribution of air quality data, integral for subsequent analyses. The data was then grouped by week, day, and hour to structure the temporal aspects. This grouped data transitioned into a GeoDataFrame, marking the shift towards spatial analysis.

With the spatial data prepared, the city's graph network was generated to establish the framework for path analysis. The pathfinding process utilized a shortest path algorithm, adapted to integrate weights based on the air quality rating, thus incorporating environmental considerations into route selection.

Finally, a 'Healthy Route' was determined, representing the outcome of applying specified criteria to the pathfinding process, ensuring that the route accounted for both efficiency and environmental health factors. The process diagram is shown in figure 3.

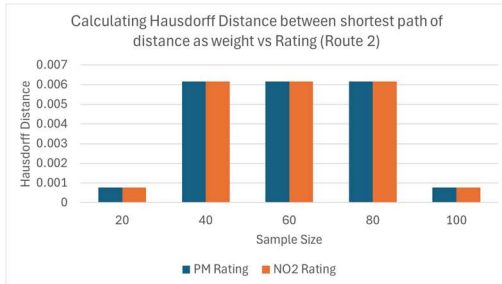


Fig. 4. Hausdorff Distance between the shortest paths of Distance as weight vs Rating (Route 2)

Sampling. We created four distinct stratified samples for each of the air quality datasets (PM2.5 and NO2), comprising 20%, 40%, 60%, and 80% of the total dataset, respectively. This methodology aimed to investigate the impact of sampling on route generation. The stratified samples were generated based on Geohash, with a precision level of 5 for PM2.5 data and a precision level of 4 for NO2 data to address data sparsity concerns.

IV. EXPERIMENTAL EVALUATION

In this section, we summarize test settings, the datasets, in addition to the baseline methods and evaluation metrics.

A. Experimental setup

Datasets. Two hyperlocal air quality datasets from New York City, covering September to December 2021, were utilized. The first dataset focuses on PM2.5 measurements, while the second centers on NO2 levels.³ OpenStreetMap provided graph information to support the implementation of the shortest path algorithm.

Deployment. We deploy our experiments on VM that runs on Google Colab with 13 GB RAM and 2 vCPU (2 Intel(R) Xeon(R) CPU @ 2.20GHz).

Evaluation metrics. We use Hausdorff distance (a.k.a. Hausdorff metric) metric, which calculates the distance between two subsets of a metric space. Hausdorff distance is a metric that calculates maximum distance for all distance values of a point of a reference trajectory from the nearest point in compared second trajectory.

B. Experimental Results and Discussion

Experiments. Our experimental setup involved creating multiple stratified samples from the air quality data to investigate the impact of data granularity on route recommendations. We implemented various scenarios where routes were generated using both PM2.5 and NO2 data, integrating air quality ratings as route weights into the modified shortest path algorithm. This systematic approach allowed us to assess and compare route efficacy under different environmental conditions and sampling strategies efficiently.

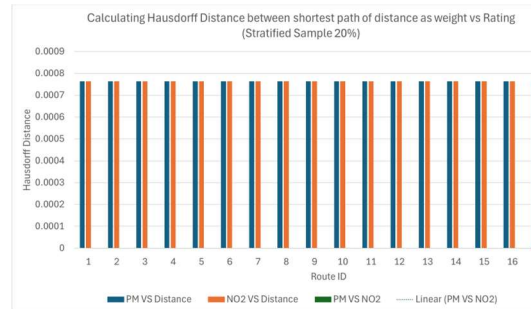


Fig. 5. Hausdorff Distance between the shortest paths of Distance as weight vs Rating (Stratified Sample 20%)

To ensure reliable and reproducible results, we conducted experiments within the Google Colab environment, utilizing the V100 High RAM hardware accelerator for enhanced computational capabilities. All experiments were coded in Python, offering a versatile and efficient framework for implementation and analysis.

² <https://www.ibm.com/docs/en/spss-modeler/saas?topic=nodes-space-time-boxes-node>

³ Dataset link: <https://zenodo.org/records/7961851>

The experimental design encompassed a series of systematic procedures aimed at evaluating the impact of sampling on route generation. Initially, we defined two distinct routes: one from Richard Morris Hunt to Hotel Belleclaire (route 1), and the other from Greenbelt Nature Center to Kids Korral (route 2). These routes were meticulously crafted using PM2.5 population data, employing varying percentages ranging from 20% to 80%. Concurrently, routes were also devised using NO2 data, mirroring the same percentage increments and utilizing both population and stratified sampling techniques.

Subsequently, the experimental scope expanded to include an additional 16 routes, supplementing the initial routes 1 and 2. These supplementary routes followed the experimental framework outlined in Table III. Notably, in this subsequent iteration, all routes, including the initial routes 1 and 2, were generated using a consistent 20% stratified sample for both PM2.5 and NO2 data. This iterative approach facilitated a comprehensive evaluation of sampling methodologies and their influence on route generation across diverse scenarios.

Our results are systematically presented in several tables and figures, highlighting the effectiveness of the modified routing algorithm. Notably, routes generated using PM rating and NO2 rating are identical in all scenarios (the Hausdorff distance is close to zero between the two routes).

TABLE I. CALCULATING HAUSDORFF DISTANCE BETWEEN SHORTEST PATH DISTANCE AS WEIGHT VS RATING - POPULATION

Calculating Hausdorff Distance between shortest path Distance as weight vs Rating - Population		
Route ID	PM2.5 Rating	NO2 Rating
1	0.000763589	0.000763589
2	0.000763589	0.000763589

Table I presents the Hausdorff distances calculated between routes generated by using distance as a weight compared to routes generated by using air quality ratings for PM and NO2 population data or 100% of the data, before sampling. The Hausdorff distance between the shortest path created using the distance as the weight seems to be the same when compared to PM rating and NO2 rating. This is also because there is no difference between the path generated using the PM rating and NO2 rating; the Hausdorff distance is 0 between them.

TABLE II. HAUSDORFF DISTANCE BETWEEN THE SHORTEST PATHS OF *DISTANCE AS WEIGHT VS RATING* (ROUTE 2)

Calculating Hausdorff Distance between shortest path of distance as weight vs Rating (Route 2)		
sample size	PM Rating	NO2 Rating
20	0.000764	0.000763589
40	0.006156	0.006156402
60	0.006156	0.006156402
80	0.006156	0.006156402
100	0.000764	0.000763589

Table II and Figure 4 shed light on the disparity in Hausdorff distance calculations when generating the same route (from Greenbelt Nature Center to Kids Korral NYC). The table illustrates the Hausdorff distance for the PM rating versus distance as weight, as well as the NO2 rating versus distance as weight. Interestingly, it becomes evident that when employing stratified sample sizes of 40%, 60%, or 80%, the distances are notably larger compared to when using a 20% sample size or population data. This observation holds true for both datasets: PM2.5 and NO2. This notable variation prompts further investigation to unravel the underlying factors contributing to this discrepancy.

The calculated Hausdorff distance between the shortest paths when weight is based on distance versus when it is based on ratings (PM rating and NO2 rating respectively) is negligible for both, roughly 0.0007635 for PM and NO2. Notably, this analysis pertains to the 20% stratified sample. Remarkably, for every route generated (totaling 18), the Hausdorff distance calculation remains consistent across both datasets. Additionally, the Hausdorff distance between routes using the PM rating and those using the NO2 rating is consistently equal to 0, indicating identical routes with no discernible difference. Figure 5 visually depicts this relationship through a bar graph, where route ID is plotted on the x-axis and Hausdorff distance on the y-axis. The graph illustrates the relationship between PM vs. distance, NO2 vs. distance, and PM vs. NO2.



Fig. 6. Optimized less polluted route (green and longer) Vs. distance-based route (red and shorter)

Figure 6 shows the two routes generated by our optimized routing engine, the green route (slightly longer distance for grayscale view of this figure)

generated using PM rating, as opposed to the red route (shorter but more polluted), generated by relying on ground distance calculation.

TABLE III. ROUTES GENERATED IN EXPERIMENTS

Routes Generated in Experiments		
Route ID	Routing from	Routing to
1	Richard Morris Hunt	Hotel Belleclaire
2	Greenbelt Nature Center	Kids Korral
3	Postcards	Greenwich Village
4	Maison sur les toits	NYSkyride
5	JFK	NBC Studios
6	Wyckoff House Museum	Eleanor Roosevelt High Residence
7	The Frederick Hotel	Rose Garden
8	The South Pole	La Casa de los Tenebaums
9	Delacorte Musical Clock	Le Carrousel
10	Friends Building	Chelsea Market
11	Grand Central Terminal	Times Square Ball
12	New York Stock Exchange	Madagascar!
13	World of Birds	Butterfly Garden
14	World of Reptiles	Congo Gorilla Forest
15	Brooklyn Bridge	Gapstow Bridge
16	Coney Island Cyclone	Central Park Carousel
17	Times Sqaure	One World Trade Center
18	Manhattan Bridge	Lower East Side Skatepark

V. LITERATURE

This survey of recent literature reveals significant contributions to urban planning through the integration of environmental data, with each paper focusing on unique challenges, methodologies, and applications related to urban environmental health.

The technical discussion transitions into a more focused examination of how spatial data is managed within smart city frameworks, as demonstrated by eight papers that delve into Sampling Design, Approximate Query Processing (AQP), and Spatial Join techniques.

[6] and [7] both concentrate on integrating spatial-aware sampling and AQP into streaming platforms like Apache Spark, but [6] emphasizes real-time quality of service enhancements by balancing latency and accuracy, while [7] aims to ensure efficient real-time query processing with a focus on urban planning and management through the SpatialSPE system. Both papers demonstrate how dynamically managed sampling can significantly improve data responsiveness and accuracy over time in urban environments.

[8] and [9] introduce systems, GeoRAP and ApproxSSPS respectively, that utilize novel sampling techniques to optimize query processing. GeoRAP in [8] employs polygon simplification and geohash-based sampling to efficiently manage spatial queries under stringent quality of service constraints. On the other hand, ApproxSSPS leverages SAOS and ex-SAOS methods to dynamically adjust to the spatial density and distribution of data, integrating these capabilities into smart city data stream management. These approaches not only enhance data throughput but also maintain high accuracy in urban analytics.

[10], [11], [12], and [13] delve deeper into the spatial join aspect, each bringing a unique perspective on processing and integrating heterogeneous urban data sets. [10] focuses on combining spatial-aware sampling designs with AQP to manage high-velocity geospatial data streams effectively, optimizing spatial joins in smart city applications. In contrast, [11] employs a novel spatial join strategy using filter-and-refine techniques combined with geohash encoding to manage mobility, pollution, and meteorological data seamlessly.

The MeteoMobil system in [12] utilizes efficient spatial join techniques, including a filter-refine method and geospatial encoding to merge mobility and meteorological data, which supports complex urban analytics. This system shows substantial reductions in processing time and computational resources. Similarly, [13] addresses spatial join processing in MongoDB, employing data partitioning and a custom spatial QoS-aware optimizer to improve query performance, which is critical for handling large and diverse datasets found in smart cities.

The varied focus across these papers—from foundational enhancements in data streaming to specific applications in spatial join processing—demonstrates a comprehensive advancement in smart city data infrastructure. Each paper contributes to a more nuanced understanding and capability in handling the complexities of big spatial data, driving forward the efficiency and responsiveness of urban systems and planning initiatives. This cohesive analysis underscores the synergy and individual strengths of these research efforts in shaping future smart city technologies.

VI. CONCLUSION & FUTURE WORKS

The routing engine developed in this study has demonstrated its effectiveness in guiding users along paths with reduced PM exposure by leveraging air quality ratings as route weights. This approach has consistently generated routes for both PM2.5 and NO2 data, underscoring its versatility and reliability.

Moving forward, future endeavors will focus on the following, web Application Integration, incorporating the routing tool into a user-friendly web application to enable real-time access by the public, expanding Geohash Encoding Techniques, exploring alternative encoding systems such as H3 and S2 to enhance route precision and accuracy, comprehensive Data Analysis, leveraging full-year air quality data to further refine and optimize the routing model for improved performance. Broader Application: Extending the methodology to cities worldwide, particularly those grappling with severe pollution issues, to validate its applicability on a global scale.

In summary, the utilization of the shortest path algorithm with air quality ratings as route weights has

resulted in identical routes generated from both PM2.5 and NO2 data, with Hausdorff distances consistently measuring at 0. These findings pave the way for the development of effective strategies in mitigating PM pollution and improving urban air quality.

REFERENCES

- [1] J. Grigg, "Particulate air pollution and vulnerability to respiratory infections in children," in *Allergens and Respiratory Pollutants*: Elsevier, 2011, pp. 213-227.
- [2] J.-P. Putaud *et al.*, "Impact of 2020 COVID-19 lockdowns on particulate air pollution across Europe," *EGUsphere*, vol. 2023, pp. 1-21, 2023.
- [3] C. E. Keenan, R. Kelleher, and S. G. Gray, "Environmental pollution, epigenetics, and cancer," in *Epigenetic Cancer Therapy*: Elsevier, 2023, pp. 175-194.
- [4] E. F. Choma, J. S. Evans, J. K. Hammitt, J. A. Gómez-Ibáñez, and J. D. Spengler, "Assessing the health impacts of electric vehicles through air pollution in the United States," *Environment International*, vol. 144, p. 106015, 2020.
- [5] S. Ilarri, R. Trillo-Lado, and L. Marrodán, "Traffic and Pollution Modelling for Air Quality Awareness: An Experience in the City of Zaragoza," *SN Computer Science*, vol. 3, no. 4, p. 281, 2022.
- [6] I. M. Al Jawarneh, P. Bellavista, L. Foschini, and R. Montanari, "Spatial-Aware Approximate Big Data Stream Processing," in *2019 IEEE Global Communications Conference (GLOBECOM)*, 2019: IEEE, pp. 1-6.
- [7] I. M. Al Jawarneh, P. Bellavista, A. Corradi, L. Foschini, and R. Montanari, "Spatially Representative Online Big Data Sampling for Smart Cities," in *2020 IEEE 25th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, 2020: IEEE, pp. 1-6.
- [8] I. M. A. Jawarneh, L. Foschini, and P. Bellavista, "Polygon Simplification for the Efficient Approximate Analytics of Georeferenced Big Data," *Sensors*, vol. 23, no. 19, p. 8178, 2023.
- [9] I. M. Al Jawarneh, P. Bellavista, A. Corradi, L. Foschini, and R. Montanari, "QoS-Aware Approximate Query Processing for Smart Cities Spatial Data Streams," *Sensors*, vol. 21, no. 12, 2021, doi: 10.3390/s21124160.
- [10] I. M. Al Jawarneh, P. Bellavista, A. Corradi, L. Foschini, and R. Montanari, "SpatialSSJP: QoS-Aware Adaptive Approximate Stream-Static Spatial Join Processor," *IEEE Transactions on Parallel and Distributed Systems*, 2023.
- [11] I. M. Al Jawarneh, L. Foschini, and P. Bellavista, "Efficient Integration of Heterogeneous Mobility-Pollution Big Data for Joint Analytics at Scale with QoS Guarantees," *Future Internet*, vol. 15, no. 8, p. 263, 2023.
- [12] I. M. Al Jawarneh, P. Bellavista, A. Corradi, L. Foschini, and R. Montanari, "Efficiently Integrating Mobility and Environment Data for Climate Change Analytics," in *2021 IEEE 26th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, 2021: IEEE, pp. 1-5.
- [13] I. M. Al Jawarneh, P. Bellavista, A. Corradi, L. Foschini, and R. Montanari, "Efficient QoS-Aware Spatial Join Processing for Scalable NoSQL Storage Frameworks," *IEEE Transactions on Network and Service Management*, 2020.