

# Breathing Green: Maximising Health and Environmental Benefits for Active Transportation Users Leveraging Large Scale Air Quality Data

Sen Yan, Shaoshu Zhu, Jaime B. Fernandez, Eric Arazo Sánchez,  
Yingqi Gu, Noel E. O'Connor, David O'Connor and Mingming Liu

**Abstract**— Pollution in urban areas can have significant adverse effects on the health and well-being of citizens, with traffic-related air pollution being a major concern in many cities. Pollutants emitted by vehicles, such as nitrogen oxides, carbon monoxide, and particulate matter, can cause respiratory and cardiovascular problems, particularly for vulnerable road users like pedestrians and cyclists. Furthermore, recent research has indicated that individuals living in more polluted areas are at a greater risk of developing chronic illnesses such as asthma, allergies, and cancer. Addressing these problems is crucial to protecting public health and maximising environmental benefits. In this project, we explore the feasibility of tackling this challenge by leveraging big data analysis and data-driven methods. Specifically, we investigate the recently released Google Air Quality dataset and devise an optimisation strategy to suggest green travel routes for different types of active transportation users in Dublin. To demonstrate our achievement, we have developed a prototype and have shown that citizens who use our model to plan their outdoor activities can benefit notably, with a significant decrease of 17.87% on average in pollutant intake, from the environmental advantages it offers.

**Index Terms**—Air Pollution, Data Analysis, Micromobility, Recommendation System, Route Planning

## I. INTRODUCTION

The detrimental effects of air pollution, specifically nitrogen oxides ( $\text{NO}_x$ ) and particulate matter (PM), on human health have emerged as a significant and escalating concern. These pollutants are predominantly emitted from diverse sources such as vehicle exhaust, industrial activities, and residential heating. The well-documented consequences on human health are substantial, resulting in a considerable number of excess deaths globally (7 million) and in Ireland (3300) annually.

$\text{NO}_x$  refers to a group of gases that includes nitrogen oxide (NO) and nitrogen dioxide ( $\text{NO}_2$ ), both of which are significant contributors to air pollution. These gases have the potential to irritate the respiratory system and worsen pre-existing respiratory conditions, such as asthma, which is a prevalent health issue in Ireland (ranking fourth highest in the world). They can also increase the susceptibility to respiratory infections. Furthermore, long-term exposure to  $\text{NO}_x$  has been associated with the development of cardiovascular diseases, including heart attacks and strokes. The

S. Yan, S. Zhu, J.B. Fernandez, E. Arazo, M. Liu and N. E. O'Connor are with the School of Electronic Engineering and SFI Insight Centre for Data Analytics at Dublin City University, Ireland. Y. Gu is with Alkermes plc, Dublin, Ireland. D. O'Connor is with the School of Chemical Sciences at Dublin City University. This project is now available on GitHub and can be accessed via <https://github.com/SFIEssential/BreathGreen>. Corresponding author's email: mingming.liu@dcu.ie

elevated concentrations of  $\text{NO}_x$  in urban areas, particularly in proximity to busy roads and industrial zones, pose a greater health risk to individuals, especially vulnerable groups such as the young, elderly, and those with preexisting health conditions.

Particulate matter (PM) refers to small particles suspended in the air, which are classified into two main categories based on their size:  $\text{PM}_{10}$  (particles with a diameter of 10 micrometres or less) and  $\text{PM}_{2.5}$  (particles with a diameter of 2.5 micrometres or less). These particles have the ability to penetrate deeply into the respiratory system when inhaled, leading to respiratory issues such as coughing, wheezing, and shortness of breath. Moreover, they can have detrimental effects on cardiovascular health, including an increased risk of heart attacks and irregular heartbeats. Additionally, exposure to PM has been associated with the development of lung cancer and adverse birth outcomes.

In Ireland,  $\text{NO}_x$  and PM emissions originate from various sources. Road transportation, especially vehicles powered by diesel engines, constitutes a significant contributor to  $\text{NO}_x$  emissions. The combustion of solid fuels, such as coal and peat, for residential heating purposes, is another prominent source of PM emissions. Industrial activities encompassing power generation and manufacturing processes also contribute significantly to the release of both  $\text{NO}_x$  and PM pollutants into the atmosphere. It is crucial to address these diverse emission sources and implement effective measures to mitigate their impact on air quality and public health.

Mitigating exposure to air pollution can have significant benefits within Ireland, as the health concerns associated with it impose a cost of approximately €2.3 billion on the Irish State annually. Studies focusing on the environmental and public health implications in the transportation sector have underscored the urgent need to reduce greenhouse gas emissions and suspended particulate matter. Notably, research has been conducted to assess the levels of traffic-related pollutants on various school commuting routes in Canada [1] and China [2], emphasising the importance of mitigating pollutant emissions, developing smart mobility devices that prioritise user health, and identifying more efficient routes with improved air quality. These efforts are crucial for safeguarding public health, especially for vulnerable road users.

In our research, we placed emphasis on identifying routes with minimal air pollution levels and offering corresponding recommendations to transportation users. The primary contr-

ibutions of our work can be summarised as follows:

- The analysis has been conducted on the air quality data from May 2021 to August 2022 in Dublin, Ireland.
- Real data has been processed, which included filling missing data values, based on geographic relevance and used in experiments to examine the performance of our system in different situations.
- A system has been developed to recommend routes to minimise pollutant inhalation for active transportation users.

The structure of this paper is as follows. In Section II, we provide a review of the relevant literature on the environmental and public health implications in the transportation field. The research question is explained in detail in Section III. Section IV describes the system architecture in our work, and Section V describes and analyses our algorithm, the dataset and the experiments. The relevant discussions and analysis are included in Section VI. Finally, we conclude our work in Section VII and discuss future plans and potential improvements.

## II. RELATED WORKS

Different approaches have been employed to address the challenge of reducing pollutant emissions, particularly in the context of route planning with an emission cap for green vehicles or automated vehicles. This topic has gained significant attention in recent research [3]–[6]. On the other hand, studies have also explored the development of smart mobility devices aimed at minimising the impact of polluting devices on individuals. For instance, in [7], a cyber-physical control system was designed to regulate the interaction between cyclists and electric motors. This system effectively manages the ventilation rate of cyclists by intelligently employing electric motor assistance. Similarly, an algorithm was developed in [8] to control the switching of hybrid vehicle motors between electric and combustion modes, aiming to protect cyclists from the harmful effects of exhaust-gas emissions. Other similar studies have also been conducted in [9]–[11].

Moreover, researchers have also explored methodologies for user-health-friendly path planning. In the study mentioned in [12], the authors propose the concept of personalised commute optimisation as a means to significantly reduce the risk of pollution exposure. They highlight the potential benefits of tailoring commute routes and modes of transportation to individual users, taking into account factors such as pollution levels, personal health conditions, and preferences. Additionally, the authors present a proof-of-concept system for optimising pollution inhalation during the commute. This system aims to provide users with personalised recommendations for travel routes and transportation options that minimize their exposure to pollutants, thereby promoting healthier and more sustainable commuting practices.

A recent study [13] investigated road safety and exposure to air pollution along three routes from central Edinburgh to the science and engineering campus of the University of Edinburgh. The results revealed significant differences in ultrafine particle exposure across the three routes. Additionally,

a study conducted in Celje, Slovenia, proposed an alternative route to reduce cyclist exposure to black carbon based on the analysis of air quality data [14]. In the study referenced [15], a cyber-physical system is proposed as a solution to mitigate the impact of urban pollution on cyclists. The system operates by indirectly controlling the breathing rate of cyclists when they traverse polluted areas. A notable study conducted in [16] focused on optimising the routing plan for smart city users with a specific emphasis on reducing  $PM_{10}$  levels. This was accomplished by implementing the Empirical Bayesian Kriging (EBK) model during the geospatial analysis phase. In a subsequent work [17], the authors extended their approach by incorporating convolutional neural networks to extract features and utilising long short-term memory models for capturing the sequential dependency of air quality. The results consistently demonstrated that although the predicted path might be longer than the shortest route, it effectively minimises the risk of pollution exposure.

Furthermore, the authors devised an intelligent speed advisory system in [18] that provides recommendations on a suitable speed for a group of cyclists, considering the varying fitness levels within the group or the different levels of electric assistance, particularly in cases where some or all cyclists utilize electric bikes. Similarly, the authors [19] developed a random forest land-use model to map  $PM_{2.5}$  concentrations by investigating the exposure among cyclists in three Asian cities. By employing spatial k-fold cross-validation, the authors found that the proposed model achieved an  $R^2$  value exceeding 0.67. Notably, the study concluded that for cycling commuters, selecting alternative routes could result in a reduction of over 30% in  $PM_{2.5}$  exposure.

However, the studies mentioned earlier could benefit from the utilisation of real-world data and the consideration of multiple pollutants. In our work, we address these limitations by employing an optimisation algorithm to identify the route with the minimum air pollution levels. This is achieved by utilising real-world data that encompasses multiple pollutants, allowing users to customise the weights assigned to each pollutant based on their preferences and requirements.

## III. RESEARCH PROBLEM STATEMENT

In this section, we describe the system model for our application. The main objective of our application is to recommend a green route to a cyclist that minimizes their exposure to a specified pollutant index while travelling from a specific starting point to a destination within a bounded geographical area. Specifically, we model the area of interest using a weighted directed graph  $G = (V, E, W)$ , where  $V = \{1, \dots, n\}$  is the vertex set, and  $E = \{e_1, \dots, e_m\}$  is the set of edges with corresponding positive weights in  $W = \{w_1, \dots, w_m\}$ . The nodes in the graph represent intersections or points of interest, while the edges indicate the road segments that connect these nodes. The weight  $w_i$  associated with the edge  $e_i$  represents the aggregated level of pollutant exposure that a cyclist may encounter while travelling along that edge, based on historical data

measurements collected at that edge and the vulnerability of the cyclist to each pollutant being considered. To further illustrate this point, mathematically, let  $K$  denote the total number of pollutants being considered, and let  $p_{ij}$  be the average pollutant level for the  $j$ 'th pollutant measured at the  $i$ 'th edge, where  $i = 1, \dots, n$  and  $j = 1, \dots, K$ , and  $l_i$  be the length of the  $i$ 'th edge, then  $w_i$  can be defined as follows:

$$w_i = l_i \sum_{j=1}^K \alpha_j p_{ij} \quad (1)$$

where  $\alpha_1, \alpha_2, \dots, \alpha_K$  are positive coefficients capturing the vulnerability of each pollutant to the cyclist, subject to the constraint  $\sum_{j=1}^K \alpha_j = 1$  for normalization. It is worth noting that given the diversity of different pollutants when they are combined for impact assessment, a min-max scaling is employed for each type of pollutant, implying that  $p_{ij} \in [0, 1]$  and  $w_i \in [0, l_i]$  which can be seen as a parameter proportionally weighted based on the length of the  $i$ 'th edge road segment  $l_i$  and the vulnerability of the cyclist to the  $K$  types of pollutants. This further implies that if the cyclist's vulnerability across different road segments is the same, the optimization model should favour the segment with the shortest length.

Given this context, our optimization problem is to find a path  $p = \langle v_1, v_2, \dots, v_h \rangle$  which consists of a sequence of vertices starting from the source vertex  $v_1 \in V$  and terminating at the destination vertex  $v_h \in V$  that minimizes the sum of the weights of its constituent edges. Mathematically, we have:

$$\operatorname{argmin}_p \sum_{i=1}^{h-1} w(v_i, v_{i+1}) \quad (2)$$

where  $w(u, v) \in W$  denotes the weight of the edge connecting vertices  $u$  and  $v$ .

#### IV. SYSTEM ARCHITECTURE

This section introduces the proposed system architecture designed to address the research questions outlined in Section III. An overview of the system architecture is presented in Fig. 1. The diagram illustrates the components of our system, which consists of a mobile application serving as the frontend to interact with users, a backend server responsible for data transfer and processing, a database for data storage, and an optimisation algorithm to generate optimised routes based on input road parameters. Our project is open-source on GitHub<sup>1</sup>, so any interested readers are encouraged to visit and check out.

##### A. Frontend

Our Android mobile application was developed using IntelliJ IDEA and subsequently tested on an Android Virtual Device featuring a *Google Play Intel Atom (x86)* processor. The interactive map functionality within the application was

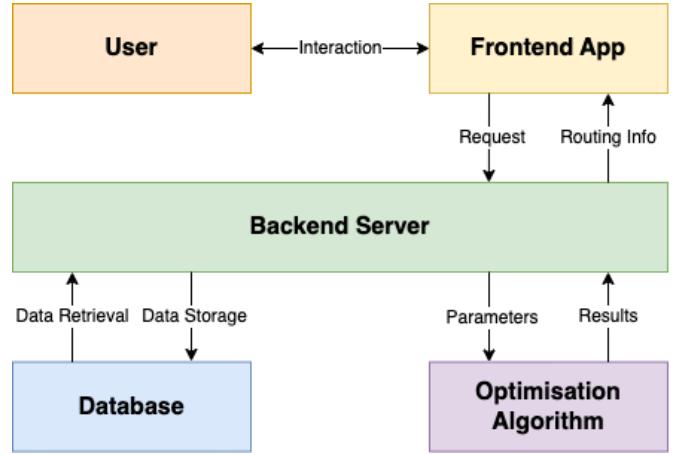


Fig. 1. An overview of the system architecture

implemented using the Google Maps API<sup>2</sup>, which enables the display of origin and destination locations, as well as routes with the shortest distance or lowest pollutant index.

##### B. Backend

Based on a data-driven approach, for the purpose of real-time response, we utilised Flask [20], a backend framework written in Python. Following the design of REST, the Flask server connects with frontend using JSON and calculates various metrics for self-optimisation. When starting the server, it loads the geographical information of Dublin from OpenStreetMap and appends the pollutant data collected from Google to generate a new digital map containing pollutant attributes on the top of roads. The routing algorithm that we propose calculates pollution level based on different setting acquired from users and output the entire route from start to the end with lowest degree of pollutants around.

##### C. Database & Algorithm

The interaction between our database and the backend is bidirectional. Data received from the backend is stored in our database, and the database is also used to retrieve and return data to the backend as per its requirements.

Similarly, the optimisation algorithm generates path recommendations based on the parameters received from the backend. The algorithm processes this information, calculate the optimised path and returns the results to the server. A more detailed explanation of the algorithm will be provided in the subsequent section.

#### V. ALGORITHM & EXPERIMENTS

##### A. Algorithm

There are several well-known algorithms available to address (2), including Dijkstra's algorithm, Floyd-Warshall algorithm, and Bellman-Ford algorithm [21]. These algorithms are designed to identify the shortest or most efficient path between two points in a weighted directed graph. For ease of reference, we shall now present a description of Dijkstra's Algorithm with reference to [22], [23] below. In Algorithm 1,

<sup>1</sup><https://github.com/SFIEssential/BreathGreen>

<sup>2</sup><https://developers.google.com/maps/documentation/directions>

the vertex set is denoted by  $Q$ , and the array  $dist$  contains the current distances from the starting point  $s$  to other vertices, where  $dist[v]$  represents the current distance from  $s$  to vertex  $u$ . The array  $prev$  contains pointers to the previous-hop nodes on the shortest path from  $s$  to the given vertex, or equivalently, it is the next hop on the path from the given vertex to the source. The variable  $alt$  denotes the length of the path from the root node to the neighbour node  $v$  if it were to go through  $u$ . If this path is shorter than the current shortest path recorded for  $v$ , the current path would be replaced with this  $alt$  path.

---

**Algorithm 1** Dijkstra's Algorithm

---

**Input:**  $G = (V, E, W)$ , a weighted directed graph, and  $s, d \in V$ , the starting point  $s$  and destination point  $d$

**Output:**  $p$ , the shortest path for (2)

```

for all vertex  $v$  in graph  $G$  do
     $dist[v] \leftarrow infinite$ 
     $prev[v] \leftarrow undefined$ 
    Add vertex  $v$  to vertex set  $Q$ 
end for
 $dist[s] \leftarrow 0$ 
while  $Q$  is not empty do
     $u \leftarrow argmin_i\{dist[i] | \forall i \in Q\}$ 
    Remove  $u$  from  $Q$ 
    for all neighbour  $v$  of  $u \in Q$  do
         $alt \leftarrow dist[u] + w(u, v)$ 
        if  $alt < dist[v]$  then
             $dist[v] \leftarrow alt$ 
             $prev[v] \leftarrow u$ 
        end if
    end for
end while
Set  $p$  as the shortest path to  $d$  using  $prev$  array

```

---

## B. Dataset

This section describes the `AirView_Dublin_City` (ADC) dataset which is recently released and the preprocessing steps followed to address challenges faced during the data collection stage. Additionally, we provide insights and visualizations of the resulting dataset.

**Google Project Air View Data - Dublin City (May 2021 - August 2022):** This dataset was collected by Google and Dublin City Council as part of Project Air View Dublin. Google's first electric Street View car, equipped with Aclima's mobile air sensing platform, drove through the roads of Dublin City measuring street-by-street air quality. The data collection process predominantly took place from Monday to Friday between 9:00 am and 5:00 pm from May 2021 to August 2022. Consequently, the dataset represents typical daytime weekday air quality. The sensors on the car measured pollution on each street and highway at 1-second intervals while driving with the flow of traffic at normal speeds. The pollutants included in the dataset are Carbon Monoxide (CO), Carbon Dioxide (CO<sub>2</sub>), Nitrogen Dioxide (NO<sub>2</sub>), Nitric Oxide (NO), Ozone (O<sub>3</sub>), and Particulate

Matter PM<sub>2.5</sub> (including size-resolved particle counts from 0.3 - 2.5  $\mu m$ ). The following two versions of the dataset are publicly available<sup>3</sup>:

- `Airview_Dublin_City_Measurements` (ADC-M) is the 1-second intervals data captured during the period.
- `AirView_Dublin_City_RoadData` (ADC-R) is the 1-second data points aggregated in approximately 50m road segments.

In this work, the ADC-R was used because it groups the individual points/measurements from ADC-M into segments. To report the concentration of each pollutant per segment, first, they get the mean of the number of measurements on this road segment, then count the number of drive passes on this road segment. Finally, for each pollutant, they report the concentration (median of drive pass mean) in  $\mu g/m^3$ . ADC-R also contains the OSM\_IDs segments to associate this dataset with the OpenStreetMap (OSM) dataset. The OSM\_IDs nodes were extracted and fed into the optimization algorithm.

## C. Data Cleaning & Prepossessing

In this work, the ADC-R dataset was used as mentioned previously. Even though this dataset is an aggregation of the raw data ADC-M and was already processed by Google, this work required the following pre-processing:

- Data interpolation for missing and negative values: the dataset includes NaNs and negative values, which may arise from sensor malfunctions or miscalibrations, potentially impacting the optimization algorithm. To address this issue, a linear interpolation method was applied to each pollutant column. Initially, the dataset was sorted by OSMID, and all negative values were replaced with NaNs. Subsequently, a linear interpolation was performed using the Python Pandas library, enabling the estimation of missing values and the smoothing of the dataset.
- Cross-referencing of the OSM and ADC-R datasets: the optimisation algorithm utilizes the nodes and segments from the OSM dataset to determine the greener path. To establish a link between these two datasets, the OSM\_IDs are used as a reference. It is important to note that the vehicle used to collect the Google Air Quality dataset did not traverse all streets and roads in Dublin, resulting in measurements being available only for specific segments. To assign these measurements to the corresponding segments in the OSM dataset, we conducted the matching process by utilising the OSM\_ID column to perform a matching process between the road dataset and the OSM. Segments existing in both datasets were selected and their corresponding pollutant measurements were merged. This resulted in the creation of a final dataset with the following format: [osm\_id, NO2\_ugm3, NO\_ugm3, CO2\_mgm3, CO\_mgm3, O3\_ugm3, PM25\_ugm3]. The optimization algorithm is applied to this final dataset.

<sup>3</sup><https://data.smartdublin.ie/dataset/google-airview-data-dublin-city>

- Furthermore, in order to address the issue of missing data on particular segments, a clustering algorithm is employed to identify the nearest neighbours from those segments. In particular, the position of a road segment is represented by its edge points, encompassing two endpoints and, if present, all inflexion points. Subsequently, the BallTree algorithm is applied to select the  $n$  closest neighbours for each target, which refers to the edge point in a road segment with missing data, within our dataset. The attributes of the target are then determined by calculating the mean value of its neighbours' attributes, thereby enabling the missing data to be filled in based on geographic relevance. The histogram depicted in Fig. 2 provides a visualisation of the distances between edge points with missing data and their closest neighbours. Considering the observed distribution of distances, the selection of 3, 5, or 10 neighbours is made. With a distance threshold of 3 km being assigned to limit the longest distance between a specific target to its neighbours, the fraction of the interpolated segments with their value assigned using the 3, 5 and 10 closest neighbours are 34.10%, 30.84% and 23.95% respectively. In this study, we selected 5 as the neighbour number because it provides a good filling rate with an acceptable number of supportive neighbours for each road segment.
- Finally, we delimited the area of interest with a focus on the central region of Dublin. In order to cover the areas frequented by shared bikes, we define the area of interest based on the map of NOW dublinbikes<sup>4</sup> and all segments and nodes within the coordinates [-6.230852, 53.359967, -6.310015, 53.330091] (as shown in Fig. 3) were extracted.

The data before and after cleaning and preprocessing is visually depicted on the map in Fig. 4, where the orange boundary corresponds to the area shown in Fig. 3 and contains 13,697 road segments in total. The blue routes on the map, with a number of 7,710, represent the road segments with Google interpolations in the original dataset, the green routes, with a number of 4,224, represent the roads interpolated by our filling process and the red routes, with a number of 1,763, are the roads still with missing values even after the filling process. In one word, the data missing rate has decreased from 43.71% to 12.87% due to our data-filling process.

#### D. Green Route Planning

Our experiments were designed to assess the impact of our algorithm on the selection of green routes in various real-world scenarios. To achieve this, we generated and compared the shortest routes with the greenest routes under different setups, as outlined below:

- 1) In the first setup, we considered all pollutants and assigned them equal weights. We conducted the experiments with a thousand randomly selected origin-destination pairs on the map, focusing primarily on

walking and private micromobility modes such as private bicycles and e-bikes. Thus, we denote this setup as “**Private Device**”.

- 2) In the second setup, we again considered all pollutants and assigned them equal weights. However, this time we conducted the experiments with ten origin-destination pairs, selecting GPS coordinates from the map of NOW dublinbikes. This setup specifically targeted transportation systems with shared devices, such as shared bicycles and shared e-bikes, and we call this setup “**Shared Device**”.
- 3) In the third setup, we examined the individual influence of different pollutants. We conducted the experiments with one origin-destination pair selected on the map, varying the weights assigned to different pollutants. For example, we assigned a weight of 1 to one pollutant, such as CO, and a weight of 0 to the other pollutants. This allowed us to obtain green routes that focused solely on reducing the specific pollutant. This setup aimed to evaluate the performance of our optimization algorithm under different pollutant weightings. This setup is named “**Single Pollutant**”.

By conducting these experiments, we aimed to gain insights into the performance and effectiveness of our optimization algorithm in different real-world scenarios and pollutant weightings. The results and relevant discussions are included in Section VI.

## VI. RESULTS & DISCUSSION

In this section, we present a summary of our experimental results and provide visual representations. Furthermore, we offer detailed discussions and analysis of the obtained results.

### A. Private Device

We performed a series of experiments with a thousand different origin-destination pairs, and the summarised results are presented in Table I. The findings clearly indicate that while the green path shows an average increase in trip distance of 15.15% compared to the shortest route, it simultaneously leads to a reduction in pollutant intake by 17.87%. Fig. 5 visualised the relationship between the pollutant saving rate and distance increase rate, with yellow scatter points representing the data. The graph is partitioned into six regions based on the linear boundaries defined by the equations  $y = 0$ ,  $y = 0.25x$ ,  $y = 0.5x$ ,  $y = x$ ,  $y = 2x$ ,  $y = 4x$  and  $x = 0$ . Additionally, the percentage of scatter points within each region is presented on the graph. The analysis of the graph reveals that the majority of scatter points are located in the region bounded by  $y = x$  and  $y = 2x$ . This observation indicates that, in most cases, the rate of pollutant saving exceeds the rate of distance increase. Hence, our algorithm demonstrates a favourable trade-off between reducing pollutant inhalation and the increase in trip distance.

It is worth noting that in some cases, certain routes in our experiments coincide with both the shortest path and the green path, resulting in zero values for pollutant savings and additional distance.

<sup>4</sup><https://www.dublinbikes.ie/en/mapping>

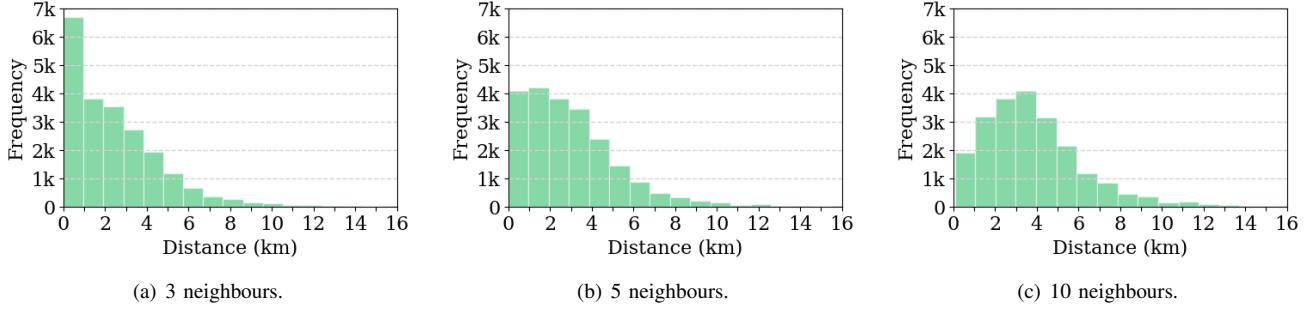


Fig. 2. Histogram of the distances to the neighbour segments used to fill the missing values for three cases: 3, 5, and 10 neighbours.

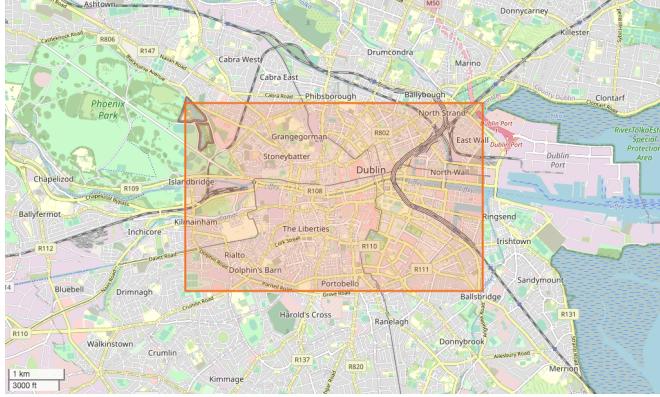


Fig. 3. Dublin area of interest

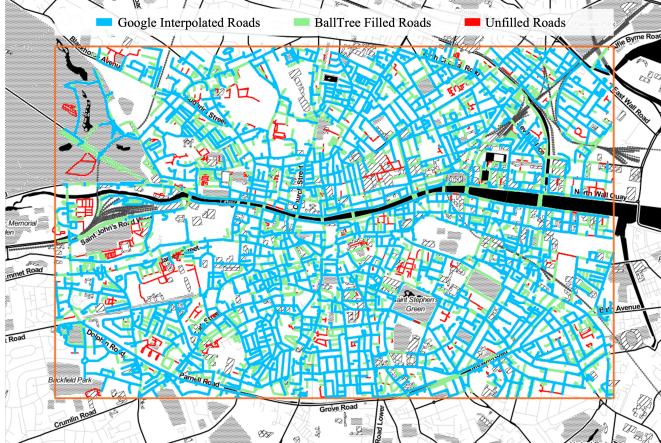


Fig. 4. Data before and after cleaning and preprocessing.

#### B. Shared Device

Another series of experiments were conducted with ten different origin-destination pairs, and the results are summarised in Table II. Similar findings can be observed in this set of experiments compared to the previous ones. Our algorithm, when applied to shared micro-mobilities, also leads to reduced pollutant inhalation at the expense of a certain increase in trip distance.

#### C. Single Pollutant

The paths for different pollutants are depicted in Fig. 6(a), with the red path indicating the shortest route, and the yellow, green, and blue paths determined by NO, CO, and NO<sub>2</sub>,

TABLE I  
PRIVATE DEVICE RESULTS.

|      | ShortPath distance (m) | GreenPath distance (m) | Pollutant saving (%) | Extra distance (%) |
|------|------------------------|------------------------|----------------------|--------------------|
| mean | 2870.31                | 3328.76                | 17.87                | 15.15              |
| std  | 1267.07                | 1572.03                | 12.20                | 13.93              |
| min  | 208.20                 | 208.20                 | 0                    | 0                  |
| 25%  | 1927.26                | 2137.51                | 7.72                 | 5.09               |
| 50%  | 2820.44                | 3284.45                | 17.20                | 12.11              |
| 75%  | 3711.74                | 4304.19                | 26.43                | 21.43              |
| max  | 7303.63                | 9265.65                | 64.99                | 83.22              |

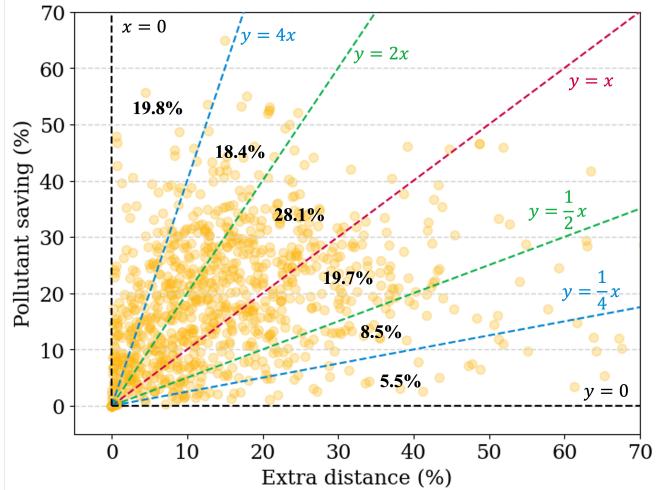
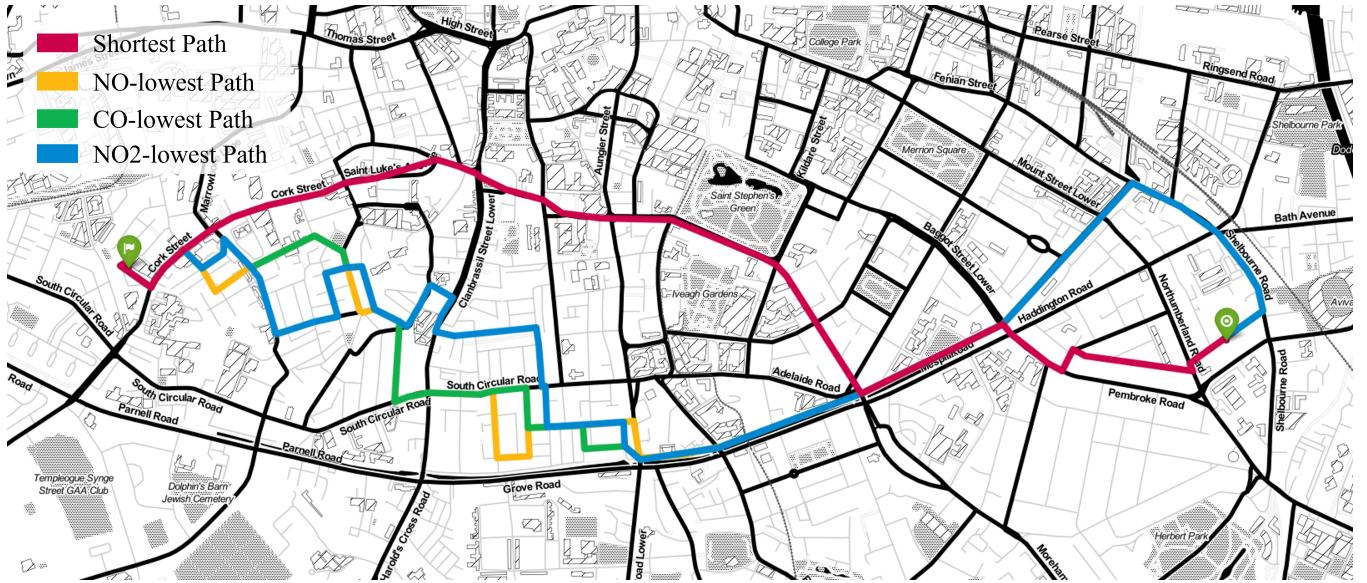


Fig. 5. Pollutant saving and distance increase.

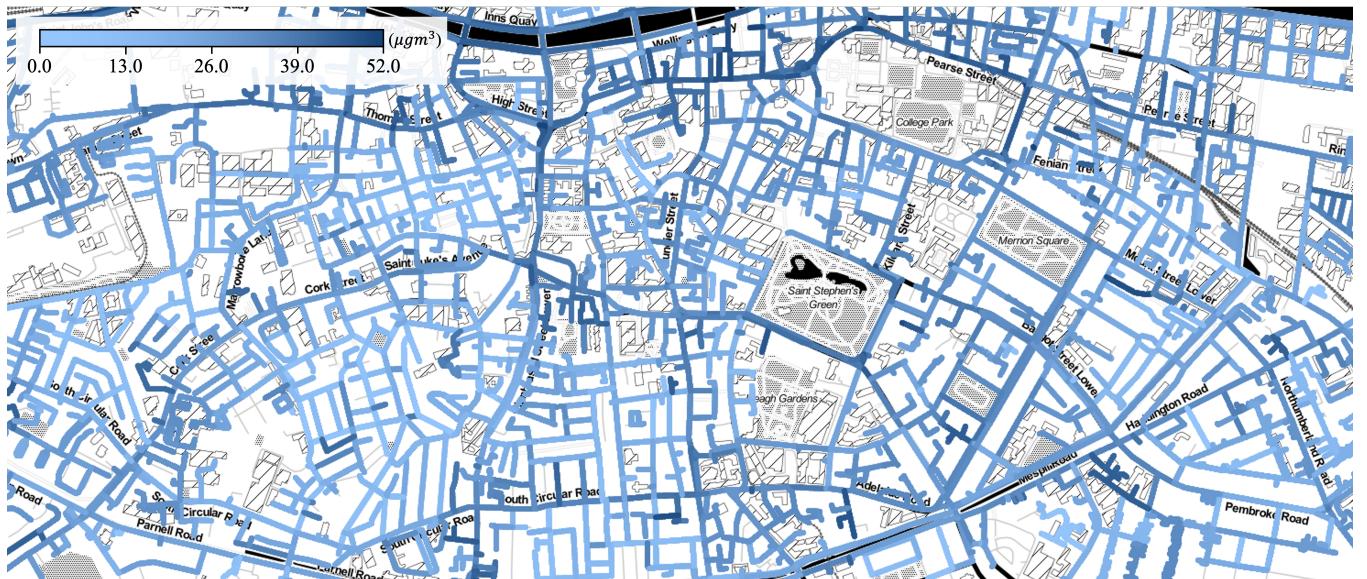
respectively, resulting in pollutant reduction percentages of 52.07%, 27.52%, and 50.57%. The values of each pollutant for different paths, as presented in Table III, demonstrate the algorithm's capability to recommend diverse routes for the same origin-destination pair based on varying pollutant weights, effectively reducing pollutant levels. For instance, taking NO<sub>2</sub> as an example, the segment values around the trip trajectory, visualized in Fig. 6(b), indicate that the blue path in Fig. 6(a) exhibits the lowest NO<sub>2</sub> levels.

## VII. CONCLUSIONS

Our research aimed to maximise health and environmental benefits for active transportation participants. We developed a system that recommends trip routes to minimise pollu-



(a) Paths determined by single pollutant.



(b) NO<sub>2</sub> distribution.

Fig. 6. Single pollutant results.

TABLE II  
SHARED DEVICE RESULTS.

|      | ShortPath<br>distance (m) | GreenPath<br>distance (m) | Pollutant<br>saving (%) | Extra<br>distance (%) |
|------|---------------------------|---------------------------|-------------------------|-----------------------|
| mean | 2460.99                   | 2921.88                   | 14.68                   | 17.59                 |
| std  | 1206.03                   | 1459.44                   | 10.37                   | 17.21                 |
| min  | 769.17                    | 769.17                    | 0                       | 0                     |
| 25%  | 1480.41                   | 1652.79                   | 5.93                    | 4.94                  |
| 50%  | 2413.75                   | 2762.74                   | 15.42                   | 15.62                 |
| 75%  | 3464.61                   | 4253.95                   | 20.84                   | 21.81                 |
| max  | 4264.24                   | 5079.85                   | 31.12                   | 54.36                 |

tant inhalation based on the analysis of air quality data in Dublin, Ireland, which was recently released by Google. Through various experiments, we evaluated the performance

of our optimisation algorithm and found it to be reliable and practical in diverse real-world scenarios, catering to both private and shared micro-mobility users, such as e-bike riders and shared bike users. Our results lead to a remarkable reduction in pollutant intake, with a significant decrease of 17.87% on average and a maximum of around 64.99%. These findings underscore the efficacy and value of our approach in improving the environmental well-being of individuals.

Furthermore, the enhancement of our work can be achieved through the implementation of a more sophisticated and rigorous pollutant evaluation standard or metrics, as well as by augmenting the completeness and comprehensiveness of the dataset. Additional avenues for improvement include conducting comparisons of pollution savings against estimated variations, such as day-to-day differences, peak and off-peak measurements, or disparities between routes in

TABLE III  
SINGLE POLLUTANT DESCRIPTIVE STATISTICS.

|                                | <b>length_m</b> | <b>CO_mgm3</b> | <b>NO_ugm3</b> | <b>NO2_ugm3</b> | <b>O3_ugm3</b> | <b>PM25_ugm3</b> |
|--------------------------------|-----------------|----------------|----------------|-----------------|----------------|------------------|
| Shortest path                  | 4556.73         | 34.16          | 1648.08        | 1585.60         | 3900.73        | 647.14           |
| CO-lowest path                 | 5643.65         | 26.88          | 853.61         | 904.94          | 3680.48        | 485.35           |
| NO-lowest path                 | 5781.24         | 31.16          | 790.19         | 951.88          | 4358.15        | 557.70           |
| NO <sub>2</sub> -lowest path   | 6127.13         | 32.30          | 1055.97        | 786.45          | 4573.79        | 588.11           |
| O <sub>3</sub> -lowest path    | 5572.22         | 27.14          | 1053.94        | 1039.73         | 3456.92        | 475.91           |
| PM <sub>2.5</sub> -lowest path | 5572.31         | 27.16          | 1009.75        | 989.46          | 3566.17        | 472.32           |

close proximity. Another viable approach entails modifying the 'least-polluting' algorithm by introducing penalties for specific turns at intersections or incorporating diverse travel speeds based on the road segments traversed and the presence of street lights. Addressing these areas of refinement holds the potential to bolster the quality and reliability of our research, thereby contributing to a more advanced and comprehensive understanding of the environmental impact of outdoor activities and the effectiveness of our system.

#### ACKNOWLEDGEMENT

This research was conducted with the financial support of Science Foundation Ireland 21/FFP-P/10266 and 12/RC/2289\_P2 at Insight the SFI Research Centre for Data Analytics at Dublin City University. For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

#### REFERENCES

- [1] S. Elford and M. D. Adams, "Exposure to ultrafine particulate air pollution in the school commute: Examining low-dose route optimization with terrain-enforced dosage modelling," *Environmental Research*, vol. 178, p. 108674, Nov. 2019.
- [2] F. An, J. Liu, W. Lu, and D. Jareemit, "Comparison of exposure to traffic-related pollutants on different commuting routes to a primary school in jinan, china," *Environmental Science and Pollution Research*, vol. 29, pp. 43319–43340, Jan. 2022.
- [3] M. A. Islam and Y. Gajpal, "Optimization of conventional and green vehicles composition under carbon emission cap," *Sustainability*, vol. 13, p. 6940, June 2021.
- [4] H. Abdullahi, L. Reyes-Rubiano, D. Ouelhadj, J. Faulin, and A. A. Juan, "Modelling and multi-criteria analysis of the sustainability dimensions for the green vehicle routing problem," *European Journal of Operational Research*, vol. 292, pp. 143–154, July 2021.
- [5] B. Olgun, Ç. Koç, and F. Altiparmak, "A hyper heuristic for the green vehicle routing problem with simultaneous pickup and delivery," *Computers & Industrial Engineering*, vol. 153, p. 107010, Mar. 2021.
- [6] L. Cai, W. Lv, L. Xiao, and Z. Xu, "Total carbon emissions minimization in connected and automated vehicle routing problem with speed variables," *Expert Systems with Applications*, vol. 165, p. 113910, Mar. 2021.
- [7] S. Sweeney, R. Ordonez-Hurtado, F. Pilla, G. Russo, D. Timoney, and R. Shorten, "A context-aware e-bike system to reduce pollution inhalation while cycling," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, pp. 704–715, Feb. 2019.
- [8] A. Herrmann, M. Liu, F. Pilla, and R. Shorten, "A new take on protecting cyclists in smart cities," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, pp. 3992–3999, Dec. 2018.
- [9] E. Dong, L. Zhang, and H. Shi, "Multi-objective optimization of hybrid electric vehicle powertrain," in *2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, IEEE, Mar. 2019.
- [10] W. Hong, I. Chakraborty, H. Wang, and G. Tao, "Co-optimization scheme for the powertrain and exhaust emission control system of hybrid electric vehicles using future speed prediction," *IEEE Transactions on Intelligent Vehicles*, vol. 6, pp. 533–545, Sept. 2021.
- [11] Y. Umezawa, K. Yamauchi, H. Seto, T. Immura, and T. Namerikawa, "Optimization of fuel consumption and NO<sub>x</sub> emission for mild HEV via hierarchical model predictive control," *Control Theory and Technology*, vol. 20, pp. 221–234, May 2022.
- [12] A. Langbridge, P. Ferraro, and R. Shorten, "Respiratory aware routing for active commuters," 2022.
- [13] J. Luengo-Oroz and S. Reis, "Assessment of cyclists' exposure to ultrafine particles along alternative commuting routes in edinburgh," *Atmospheric Pollution Research*, vol. 10, pp. 1148–1158, July 2019.
- [14] B. Jereb, T. Batkovič, L. Herman, G. Šipek, Š. Kovše, A. Gregorič, and G. Močnik, "Exposure to black carbon during bicycle commuting—alternative route selection," *Atmosphere*, vol. 9, p. 21, Jan. 2018.
- [15] S. Sweeney, R. Ordonez-Hurtado, F. Pilla, G. Russo, D. Timoney, and R. Shorten, "Cyberphysics, pollution mitigation, and pedelecs," *ArXiv*, 06 2017.
- [16] K. K. R. Samal, K. S. Babu, and S. K. Das, "ORS: The optimal routing solution for smart city users," in *Lecture Notes in Electrical Engineering*, pp. 177–186, Springer Singapore, 2020.
- [17] K. Samal, K. Babu, and S. Das, "Predicting the least air polluted path using the neural network approach," *ICST Transactions on Scalable Information Systems*, vol. 8, p. 170250, Oct. 2021.
- [18] Y. Gu, M. Liu, M. Souza, and R. N. Shorten, "On the design of an intelligent speed advisory system for cyclists," in *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, IEEE, Nov. 2018.
- [19] T.-G. Wu, Y.-D. Chen, B.-H. Chen, K. H. Harada, K. Lee, F. Deng, M. J. Rood, C.-C. Chen, C.-T. Tran, K.-L. Chien, T.-H. Wen, and C.-F. Wu, "Identifying low-PM2.5 exposure commuting routes for cyclists through modeling with the random forest algorithm based on low-cost sensor measurements in three asian cities," *Environmental Pollution*, vol. 294, p. 118597, Feb. 2022.
- [20] M. Grinberg, *Flask web development: developing web applications with python.*" O'Reilly Media, Inc.", 2018.
- [21] F. Sapundzhi and M. Popstolov, "Optimization algorithms for finding the shortest paths," *Bulgarian Chemical Communications*, vol. 50, no. Special Issue B, pp. 115–120, 2018.
- [22] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*. MIT Press, 3 ed., 2009.
- [23] R. Szczepanski and T. Tarczewski, "Global path planning for mobile robot based on artificial bee colony and dijkstra's algorithms," in *2021 IEEE 19th International Power Electronics and Motion Control Conference (PEMC)*, IEEE, Apr. 2021.