Sustainable Cities and Society 87 (2022) 104180

Contents lists available at ScienceDirect

Sustainable Cities and Society

journal homepage: www.elsevier.com/locate/scs

MobiWise: Eco-routing decision support leveraging the Internet of Things☆

Ana Aguiar [a,*], Paulo Fernandes [b], Andreia P. Guerreiro [c,d], Ricardo Tomás [b], João Agnelo [d], José Luís Santos [e], Filipe Araújo [d], Margarida C. Coelho [b], Carlos M. Fonseca [d], Pedro M. d'Orey [f], Miguel Luís [g], Susana Sargento [h]

[a] University of Porto, Instituto de Telecomunicações, Porto, Portugal
[b] University of Aveiro, Centro de Tecnologia Mecânica e Automação (TEMA), LASI - Intelligent Systems Associate Laboratory, Aveiro, Portugal
[c] INESC-ID, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal
[d] University of Coimbra, Centre for Informatics and Systems of the University of Coimbra (CISUC), Department of Informatics Engineering, Portugal
[e] University of Coimbra, Centre for Mathematics of the University of Coimbra (CMUC), Department of Mathematics, Coimbra, Portugal
[f] CISTER Research Centre, University of Porto, LASI - Intelligent Systems Associate Laboratory, Porto, Portugal
[g] Instituto Superior de Engenharia de Lisboa, Instituto de Telecomunicações, Aveiro, Portugal
[h] University of Aveiro, Instituto de Telecomunicações, Aveiro, Portugal

ARTICLE INFO

Keywords:
Eco-routing
Internet of Things
Multi-objective optimisation
Decision support

ABSTRACT

Eco-routing distributes traffic in cities to improve mobility sustainability. The implementation of eco-routing in real-life requires a diverse set of information, including different kinds of sensors. These sensors are often already integrated in city infrastructure, some are technologically outdated, and are often operated by multiple entities. In this work, we provide a use case-oriented system design for an eco-routing service leveraging Internet-of-Things (IoT) technologies. The methodology involves six phases: (1) defining an eco-routing use case for a vehicle fleet; (2) formulating a routing problem as a multi-objective optimisation to divert traffic at a relevant hub facility; (3) identifying data sources and processing required information; (4) proposing a microservice-based architecture leveraging IoT technologies adequate to a multi-stakeholder scenario; (5) applying a microscopic traffic simulator as a digital twin to deal with data sparsity; and (6) visually illustrating eco-routing trade-offs to support decision making. We built a proof-of-concept for a mid-sized European city. Using real data and a calibrated digital twin, we would achieve hourly total emissions reductions up to 2.1%, when applied in a car fleet composed of 5% of eco-routing vehicles. This traffic diversion would allow annual carbon dioxide and nitrogen oxides savings of 400 tons and 1.2 tons, respectively.

1. Introduction

According to the European Environmental Agency, and despite the pollutant emission reductions due to the pandemic restrictions, 95% and 89% of the urban population were exposed to levels of fine particulate matter and Nitrogen Oxides (NOx) above the World Health Organisation (WHO) guidelines in 2020 (European Environment Agency, 2020), with severe consequences for human health.

Eco-routing has been proposed to distribute traffic in cities to improve the sustainability of the mobility sector (Boriboonsomsin et al., 2014; Sun & Liu, 2015; Zhou et al., 2016). Urban transport is one major cause of pollutant emissions. According to TOMTOM, congestion in 4 large European cities represents 10 to 15% of total traffic emissions.[1] Additionally, traffic congestion has time and productivity costs. The latest report by the (Texas A& M. Transportation Institute, 2021) sets the yearly tolls in the U.S. at 8.7 billion hours and 190 billion US Dollars,

☆ **Financial disclosure**: This work was supported by UIDB/00481/2020, UIDP/00481/2020, UIDB/04234/2020, UIDP/04234/2020, UIDB/50008/2020, UIDP/50008/2020, UIDB/00324/2020, UID/CEC/00326/2020, FCT grants 2020.05106.BD and PTDC/CCICOM/31198/2017, and contract 2021.02488.CEECIND - FCT (Fundação para a Ciência e a Tecnologia); CENTRO01-0145-FEDER-022083 - Centro Portugal Regional Operational Programme (Centro2020), under the PORTUGAL 2020 Partnership Agreement, through the European Regional Development Fund; MobiWise (P2020 SAICTPAC/0011/2015), DICA-VE (POCI-01-0145-FEDER-029463), Driving2Driverless projects (POCI-01-0145FEDER-031923), and CityCatalyst (POCI-01-0247-709-FEDER-046119), co-funded by COMPETE2020, Portugal2020 - Operational Program for Competitiveness and Internationalization (POCI), European Union's ERDF (European Regional Development Fund), and by national funds (OE), through FCT/MCTES.

* Correspondence to: Rua Dr. Roberto Frias, 4200-465, Porto, Portugal.
E-mail address: ana.aguiar@fe.up.pt (A. Aguiar).

[1] https://www.tomtom.com/blog/traffic-and-travel-information/the-true-environmental-cost-of-inner-city-congestion/

https://doi.org/10.1016/j.scs.2022.104180
Received 17 May 2022; Received in revised form 16 September 2022; Accepted 16 September 2022
Available online 4 October 2022

counting also 3.5 wasted gallons of fuel and 18 million tons of excess greenhouse gas emissions. However, routing choices that minimise system travel time often require longer travelling distances, thus resulting in higher amounts of fuel used and pollutant emission levels (Ahn & Rakha, 2008). Eco-routing was first introduced by Ericsson et al. (2006) and has been applied to an urban network in Sweden to select the route with the lowest fuel consumption. The implementation of eco-routing in real life requires a diverse set of information, like traffic predictions, common routes, emission predictions, etc. This information can be obtained by collecting and processing data from different kinds of sensors, often already present in city infrastructure. Several works have explored the potential of fusing multiple data sources (Arsenio et al., 2020; Zhang et al., 2017), assuming that the data is available and formatted in a way to be fused. Although this is easily done for a specific data-based analysis, the deployment of such applications and systems operating in a city still runs against practical limitations like interoperability of different modules, programmatic data integration or handling sparsity (Arsenio et al., 2020). The data sources are more often than not operated under the responsibility of different entities. Some examples are different public transport fleets (bus, taxi, ride-sharing services), municipality road sensors often installed at different times and of different types and vendors (inductive loops, counting cameras, light barriers, etc.), or air quality sensors. This raises barriers to the deployment of eco-routing systems that assume that a single entity controls all parts of the system. Different sensors are challenging because of the need to integrate different types of data (quantities, units, formats), supplied in different formats through different methods and with different quality. Moreover, many of these sensors often use outdated or even obsolete technology – characteristics often identified as legacy – and do not allow programmatic integration in data processing workflows. Further, a broad range of sub-systems are necessary to process the obtained information, and different entities specialise in different types of algorithms, from information extraction and prediction (traffic, emissions, air quality, etc.) to the routing itself, as is mentioned by different scientific and technological domains.

This paper seeks to develop an eco-routing management system based on the multi-source traffic information and processing ecosystem of an urban environment. These aspects reflect the expected environment in cities and raise challenges in terms of system architecture and data source integration. We address an eco-routing optimisation problem from a trans-disciplinary perspective by developing a novel and generic traffic management system in a typical multi-stakeholder urban ecosystem that provides accurate decision-making information by quantifying alternative solutions using multi-objective optimisation, Internet of Things (IoT) technology, a city digital twin, and a visualisation portal. The proposed MobiWise decision support system focuses on the trade-offs for improved management of cities and society that involve optimising multiple often conflicting goals (travelled time versus pollutant emissions; global pollutant emissions versus local pollutant emissions), deals with information extracted from heterogeneous data sources in a multi-stakeholder ecosystem, addresses data sparsity through a digital twin, and proposes a portal with visualisations to support decision making by e.g., mobility managers. The main contributions of the paper are:

- Designing an eco-routing decision support system in a micro-service architecture to address the multi-stakeholder and multi-data source environment;
- Proposing and validating with an implementation the use of a digital twin to address data sparsity;
- Adopting a traffic distribution perspective, differently from individual decision making in most works providing IoT based solutions, and proposing a portal where decision makers can explore trade-offs of the Pareto front;
- Validating the proposed multi-criteria eco-routing decision support system for a sub-set of population under different operational and demand scenarios in a European mid-sized city;
- Comparing the impacts of eco-routing decisions between baseline and optimised scenarios for several metrics, like travel time, carbon dioxide (CO_2) and nitrogen oxides (NOx) emissions and traffic congestion, greenhouse gases (GHG), NOx, noise, and road crash external costs, which are agglomerated in the sustainability indicator defined in Section 7.

This paper is organised as follows. The next section provides a review of the relevant state-of-the-art. After presenting the methodology in Section 3, we provide a blueprint for an eco-routing system by identifying stakeholders and relationships in the use case design in Section 4, where we also identify key functions and associated system components. We then design the individual components: multi-objective optimisation (Section 5) and propose a microservice architecture glued by IoT technology to integrate diverse data sources and provide interoperability in a multi-stakeholder ecosystem (Section 6). We also propose a digital twin to deal with spatial and temporal data sparsity (Section 7) and design a decision support portal (Section 8) that offers visualisations. We validate the approach by illustrating trade-offs and quantifying savings for off-peak and peak traffic hours in a mid-sized European city (Section 9). Finally, we conclude the paper by discussing limitations and open challenges in Section 10.

2. Literature review

Eco-routing. Eco-routing is a hot topic in traffic management, proposing to leverage information gathered from smart infrastructures to improve vehicle fuel efficiency both at tactical and operational levels (Sivak & Schoettle, 2012; Zeng et al., 2020). A good body of research has developed eco-routing algorithms and systems under different traffic conditions (Bandeira et al., 2018; Boriboonsomsin et al., 2012; Zhao et al., 2015). For instance, Bandeira et al. (2018) used a platform that combined empirical and microscopic-based approaches for assessing eco-routing strategies in different types of routes. Besides travel time and emission variables, they included social criteria, namely, traffic noise and several traffic conflicts on route decisions. Other authors considered new alternative fuel types, such as hybrid electric vehicles, plug-in hybrid vehicles and electric vehicles eco-routing (Li et al., 2020; Rhun et al., 2020; Shen et al., 2019; Wang et al., 2019a). The interest is also growing concerning the implementation of eco-routing using transport systems with smart traffic control treatments and connected and autonomous vehicles applications (Djavadian et al., 2020; Ma et al., 2019). Bearing in mind the impact of traffic externalities on the local population, Fernandes et al. (2019) applied an eco-indicator for eco-routing strategies jointly expressing traffic congestion, noise, GHG and NOx emissions, health impacts and road crash-related costs. Wang et al. (2019a) developed a real-time vehicle-specific eco-routing model for use in onboard navigation systems for both internal combustion engines and battery-powered electric vehicles. However, the model neglected the impact on local pollutants emissions such as NOx, particulate matter (PM) or impacts on traffic noise.

Technology for Intelligent Mobility Services. The IoT provides the capability of collecting and processing an enormous amount of data in nearly real-time, and intelligent transport systems are one major application domain (Bansal et al., 2020). This allows for a better characterisation of the environment and systems and provides the necessary infrastructure for highly adaptive services (Zhang & He, 2020).

Mobility as a Service also relies on technology platforms (Calderón & Miller, 2020), however, such platforms remain opaque and there is no concern about becoming a part of interoperable city mobility services. Melkonyan et al. (2022) proposes a collaborative decision making framework for policy design among diverse stakeholders and conceptualises the integration of various data flows, but does not provide technical solutions and does not consider the operational perspective of traffic distribution. A multitude of intelligent transport and smart mobility services have been proposed based on that data and targeting management of vehicle's individual fuel efficiency, reduction of emissions, or development of intelligent traffic control systems. However, that research focuses on the set of components and their individual performance aspects, disregarding the perspective of how to integrate, and potentially re-use them in a diverse ecosystem. Another distinguishing aspect from many other works is that this paper adopts a planning perspective, targeting a decision support system for traffic distribution and not the perspective of an individual vehicle. The next paragraphs describe representative related work, starting with IoT frameworks, followed by eco-routing applications, and ends with data science and machine learning.

Firdhous et al. (2021) introduced an IoT framework to provide environment-aware traffic management. Authors break down the framework into several layers, each one with a specific role: data acquisition, communication, data analysis and knowledge generation and information dissemination. The services provided by each layer are presented, but no details are given about how interoperability can be achieved. Information was also scarce on what concerns the implementation and validation of the proposed solution. Mondal and Rehena (2021) presents another IoT framework for road traffic congestion management, but propose a monolithic system that does not enable the flexibilty of a micro-service architecture and without data integration or interoperability considerations. The recently started European Union Horizon 2020 project URBANITE[2] will develop modules that enable making unstructured dispersed data available for improving decision-making for urban mobility, a concept that can be instantiated as a micro-service architecture, but the project has no outputs yet.

Another difference to existing work is the perspective of urban planning. Several works adopt an ego-perspective and focus on applications to provide individual route or driving behaviour recommendations. Boriboonsomsin et al. (2012) proposes an eco-routing system that considers multiple data sources and dynamic traffic information to provide single-origin single-destination routes. Orfila et al. (2015) developed an eco-driving application for Android smartphones that connects with the vehicle via onboard diagnostics (OBD) device, enabling the analysis of past actions and the prediction of upcoming events. More recently, Priya et al. (2022) proposes a in-vehicle system that recommends driver and cruise control actions that lead to reduced CO_2 emissions. These works consider an individual perspective when addressing emission reduction, failing to address the broader societal perspective of decision support for traffic distribution.

Other works focus on individual components, e.g. communications, data science and prediction methods, disregarding the data collection and interoperability or decision support perspectives. Hussain et al. (2021) presented a cognitive-based routing decision framework based on Extreme Learning Machine and Global Navigation Satellite System (GNSS) data gathered from a vehicular communication network. The authors did not describe the operational requirements for the collection of data, and the focus lies on the machine learning framework. On the other hand, Elbery and Rakha (2019) addressed the impact of vehicular network communication losses and penetration ratio on an individual eco-routing service that fuses historic and real-time data from vehicular probes. This work deals with limitations of the data collection, but adopts an individual perspective and considers a monolithic system.

Traffic prediction is a highly active area of research, with many authors using advanced machine learning methods in an attempt to better manage the flow of vehicular traffic around road networks. As one example, the work in Majumdar et al. (2021) combined the use of IoT road traffic sensors and deep learning approaches to analyse the traffic flow of a city, which is then used to predict the propagation of congestion in the near future. Although this work uses a single source of information and inductive loops registering the vehicle's average speed, it presents outputs that are not useful for traffic management strategies. Another common perspective is fusing different data sources for mobility information extraction (Pirra & Diana, 2019; Silva et al., 2019), and not on how to programmatically integrate the data sources in a multi-stakeholder ecosystem. Yet other works propose frameworks to deal with multi-source data collection and data processing and fusion challenges to build decision support systems that attract users to public transport (Guido et al., 2017) or to provide visualisations of mobility flows to different stakeholders (You et al., 2020). These systems focus on the various challenges of data collection and processing, but propose monolithic systems for integration, and have no concerns on interoperability. They also do not address traffic distribution or eco-routing, and none of them proposes a digital twin for dealing with data sparsity or prediction.

To the best of our knowledge, no work previously has addressed the problem of designing and validating a decision support system offering visualisations for understanding the trade-offs involved in jointly eco-routing various vehicles.

Research Gap. Prior articles explored the impacts of eco-routing from an individual vehicle's perspective, with perfect information available from a monolithic system. In a realistic environment, however, data sources are diverse and not uniformly distributed and necessary components of the application are provided by different entities. Furthermore, solutions that optimise one particular criterion generally produce poor results in the remaining metrics. In this paper, we show how standard IoT technology can be used as the enabling technological bond, providing means to integrate data sources and computation components from different stakeholders in a microservices architecture. This architectural approach fits our needs, by splitting functionality into self-contained loosely-coupled components that can be separately created, deployed, and managed by independent teams. Additionally, standard IoT technology also provides interoperability for diverse data sources. We also propose a digital twin to address data sparsity. Moreover, we demonstrate how multi-criteria optimisation applied to managing a fleet can provide quantitative trade-off information to decision-makers, other than classical routing optimisation approaches that focus on optimising individual routes. Finally, we propose a portal to offer visualisations of the trade-offs involved in the different solutions to inform decision-making.

3. Methodology

The core idea of the paper is to adopt a use-case-driven design methodology (Cockburn, 1997) to system design:

1. We start by designing the **eco-routing use case** for traffic distribution in an urban area (*Section* 4). Then, we decompose the use case into sub-problems in diverse research fields.
2. We **formulate the multi-objective optimisation problem** that lies at the base of the routing service (*Section* 5), and we **identify the necessary per-edge information** to feed the problem.
3. Next, we identify **raw data sources** from which the per-edge information could be inferred. A middle sized European city serves as archetype city that motivates the assumption of a mix of novel and legacy sensor data and a multiplicity of stakeholders. We then **design the Mobiwise eco-routing application as a micro-service based system** to accommodate data sources from heterogeneous stakeholders and other necessary functions, leveraging IoT technology as a middleware that provides clearly defined interfaces to enable interoperability among the various components (*Section* 6).
4. We **build a digital twin** for the archetypal middle-sized European city, to work as a traffic and emission prediction module, addressing the data sparsity problem (*Section* 7).
5. We implement a **proof-of-concept** eco-routing assessment tool integrating the previous components to facilitate the evaluation of trade-offs for different optimisation scenarios, and enable

[2] https://urbanite-project.eu/

visualisation of impacts by decision makers (*Section* 8). This tool is made available as open-source software with the publication of this article.

6. We show results on the trade-offs of eco-routing for 3 scenarios of choice, quantifying the gains in pollutants emissions and road traffic externalities obtained even when eco-routing is applied to only 5% of the vehicles (Section 9).

The following sections describe in detail each of these system design steps and provide an example of the application of the proposed methodology in a mid-sized European city, using real-world traffic data.

4. Eco-routing use case

We consider that a subset of all vehicles in the considered area follows the route recommendations. These vehicles can represent a fleet of (e.g ride sharing, taxis, high-occupancy, micromobility, same emission standards) vehicles or the percentage of individual drivers that would commit to following route recommendations. Although actual traffic distribution through optimised routing may be impaired in the real world by the limited adoption of the proposed routes, addressing this aspect falls out of the scope of this work.

For resource allocation problems, the effects of traffic management are usually best perceived for highly loaded road transport systems. This is especially true during short periods in the peak hours when many vehicles try to simultaneously reach a specific hub facility, such as business/industrial districts, university campuses, city train stations, large residential areas or outgoing commuter junctions, sports or cultural events and shopping malls (Bandeira et al., 2020; Fernandes et al., 2018; Tomás et al., 2021). These vehicles are coming from different locations, and they tend to increase levels of congestion, emissions, noise and risk of crashes not only on the roads near the destination but also on other roads of the city network. We test these traffic patterns by applying a multi-origin, single-destination approach aiming at distributing a segment of vehicles representing a small proportion of the population in a traffic network from a specific city hub. A major potential of our approach is that it accounts for the differences in traffic patterns among demand periods and hub locations, as demonstrated in the next sections. Generally speaking, the location of city-specific origins and hubs can be obtained using different data sources, for instance, origin–destination matrices, geostatistics of economic activity or population census. Such information can be complemented with sensing data to define reliable paths along the study domain.

Optimal traffic distribution relies not only on formulating and solving the optimisation problem, but also on obtaining the necessary information to make accurate predictions. Fig. 1 shows the context diagram for the eco-routing use case, identifying the system components. Sensors collect roadway data, such as the number of vehicles, speed, or noise levels. These data are sent to the *Sensor Data Collection* module that will calculate the relevant statistics from observed data and associate it with a map edge (e.g., average vehicle speed in each street). Summarised traffic data is used to provide a prediction of relevant variables for optimisation in the *Traffic and Emissions Computation module*, namely average vehicle speed and gas emissions in each street for a certain time. These data are then fed to the *Mobiwise Optimisation module* along with static information about the network (e.g., roadways ID, the corresponding length, traffic direction, etc.). This module will determine a set of Pareto-optimal solutions considering the set of objectives to be simultaneously minimised, such as the total travel time, distance, tailpipe emissions or an aggregating eco-indicator. The results show the trade-offs of different policy choices and can be seen on the *MobiWise Portal*, along with data visualisation of emission hotspot locations. The *Decision Maker* can select the objectives to consider, e.g. time versus distance, time versus CO_2 emissions, or time versus NOx emissions, and can change them at will according to policy considerations. The implementation of the chosen policy solution, i.e. its translation into actual routes served to the *Driver/Autonomous Vehicle* using the eco-routing service, is done in the *Mobiwise Optimiser module*.

5. Multi-objective optimiser

The route optimisation problem is modelled as a (multi-objective) minimum cost flow problem (Ahuja et al., 1995; Hamacher et al., 2007), whose objective functions are selected by the decision maker, e.g. total travel time and CO_2 emissions.

Let $\mathrm{N} = \{1, \dots, n\}$ be a set of n nodes, $\mathrm{A} \subseteq \mathrm{N} \times \mathrm{N}$ be a set of m edges, and $\mathrm{G} = (\mathrm{N}, \mathrm{A})$ a directed network. Let G represent a road network where the edges represent the roadways and the corresponding traffic direction, and nodes represent, for example, road intersections. Each edge has an associated set of characteristics/costs, some of which are static, such as the lane length or the number of lanes, and others depend on the traffic conditions, such as the average travel speed and the average gas emission (e.g. CO_2, carbon monoxide (CO), NOx). These costs are meant to represent the costs associated with base flow traffic, over a period.

Consider a set of v vehicles entering the network over a (small) period, T, possibly from different nodes, and with the same destination node. The goal is to determine the best route for each of these vehicles, such that the overall costs are minimised. Each of these costs is modelled as one of d (objective) functions. Because the problem is considered for a small period, the base flow traffic is assumed not to vary much and thus, is considered to be constant over the period T. This base flow also influences the capacity of each edge, i.e., the number of (additional) vehicles that can travel through it during the period of time T. The problem is thus formulated as a multi-objective minimum cost flow problem (Hamacher et al., 2007):

Problem 1 (*Multi-objective Minimum Cost Flow Problem*).

$$\min_{x \in \mathbb{N}^{n \times n}} \quad f_k(x) = \sum_{(i,j) \in \mathrm{A}} c_{ijk} x_{ij} \qquad k = 1, \dots, d \tag{1a}$$

$$\text{subject to} \quad \sum_{\{j:(i,j) \in \mathrm{A}\}} x_{ij} - \sum_{\{j:(j,i) \in \mathrm{A}\}} x_{ji} = b_i \qquad i = 1, \dots, n \tag{1b}$$

$$0 \le x_{ij} \le u_{ij} \qquad \forall (i,j) \in \mathrm{A} \tag{1c}$$

where $c_{ij} \in \mathbb{N}^d$ is the d-dimensional cost vector associated to the arc $(i,j) \in \mathrm{A}$ per unit flow, i.e., c_{ijk} represents the cost of objective k associated to the arc (i,j). The vector $b \in \mathbb{N}^n$ represents the nodes' flow balance and is assumed to be such that $\sum_{i=1}^{n} b_i = 0$. Finally, u_{ij} represents the capacity of the arc $(i,j) \in \mathrm{A}$.

Constraint (1b) is known as the mass balance constraint (Ahuja et al., 1995). A value of $b_i = 0$, where $i \in \mathrm{N}$, indicates that the number of vehicles entering the node i (from incoming edges) and the number of vehicles leaving the node (to outgoing edges) is the same. A positive value h of $b_i = h > 0$ indicates that i is a supply node, i.e., node i is the origin node of h of the vehicles going through that node. A negative value h of $b_i = h < 0$ indicates that i is a demand node, i.e., node i is the destination node of h of the vehicles going through that node. Consequently, in the considered scenario of multiple origins and a single destination, there is at least one node $i \in \mathrm{N}$ such that $b_i > 0$ (the sum of which is v), and there is exactly one node $j \in \mathrm{N}$ such that $b_j < 0$ (in this case, $b_j = -v$).

Each component x_{ij} in a solution (flow) x of Problem 1 represents how many of the v vehicles go through each arc $(i,j) \in \mathrm{A}$. Given its multi-objective formulation, there is a set of Pareto-optimal solutions to this problem. The choice of a single one of such Pareto-optimal solution is left to the *Decision Maker*.

Note that a (Pareto-optimal) flow, x, does not provide the exact route that each vehicle should follow. The translation of a flow to routes is performed in the route assignment step. The same flow can be translated into different sets of routes.

Fig. 2a shows an example of a flow associated to two cars, each one going from one of two supply nodes, A and B (see that $b_A = 1$ and $b_B = 1$), to the demand node, F ($b_F = -2$). The flow is represented by the numbers near each arc which indicate how many of the cars go

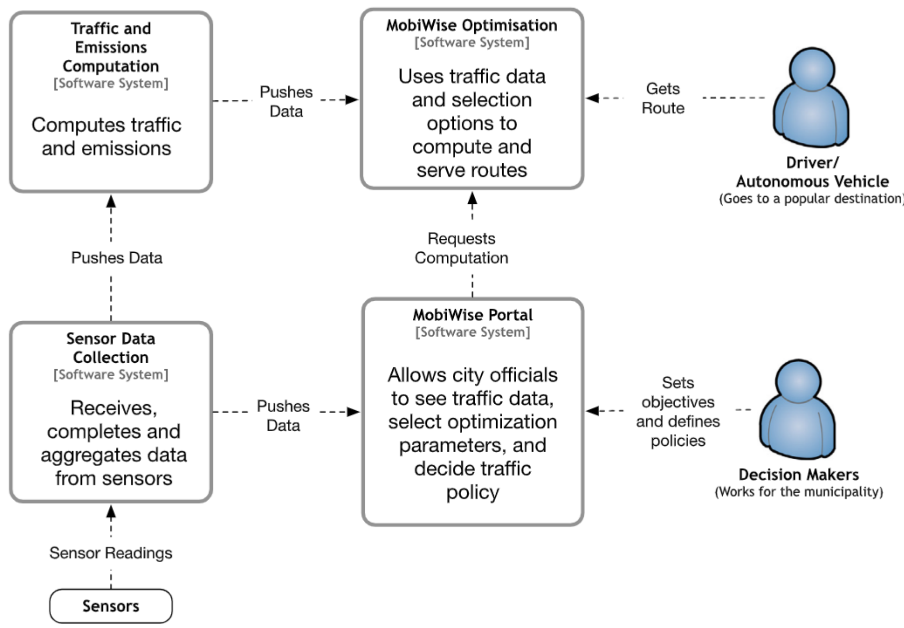

Fig. 1. Context diagram of the eco-routing use case.

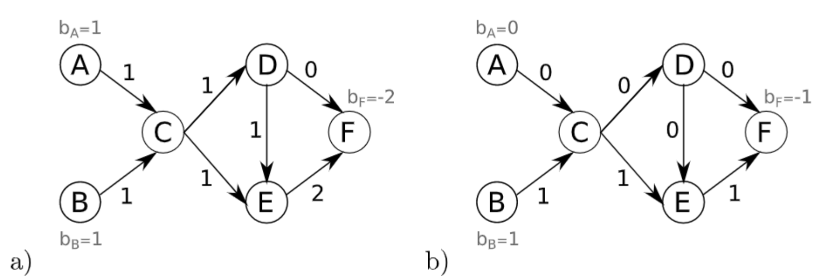

Fig. 2. Example of the translation of flow into routes.

through that arc. In this example, there are two possible translations of the flow to routes, either routes $ACDEF$ and $BCEF$ are obtained, or routes $ACEF$ and $BCDEF$.

One arbitrary way of translating a flow to routes is the following. Randomly select a supply node s, i.e., $s \in \{i \in N \mid b_i > 0\}$, and decrease the value of b_s by one. Then, select the next arc in the route by randomly choosing an arc leaving the last node and that has a positive number of flow units, i.e, an arc (s, j) such that $x_{sj} > 0$ and where s is the last node in the route being built, and decrease the value of x_{sj} by one. This step is repeated until an arc ending on the demand node is selected, in which case the route is complete and the flow left represents one less vehicle. This procedure is repeated until there is no node $s \in N$ such that $b_s > 0$, i.e., it is repeated v times to determine all v routes.

In the example of Fig. 2a, assume that the randomly selected supply node is A. Then, the only arc that can be selected is (A, C) and the flow units associated with it ($x_{AC} = 1$) are decreased by one. The next selected arc must either be (C, D) or (C, E) because both leave C and have flow units left (i.e., $x_{CD}, x_{CE} > 0$). Assuming that the arc (C, D) is selected, the next arcs selected must be (D, E) and then (E, F). Hence, the first computed route is $ACDEF$. Fig. 2b shows the remaining flow. Repeating the procedure once more will produce route $BCEF$.

6. Internet of Things as system of systems enabler

In Section 5 we identified the dynamic inputs to the optimiser as traffic and emission costs per edge for the period of concern for the optimisation. This information can be obtained through a complex process of data harvesting, processing and prediction. First, data need to be collected from multiple sources—data harvesting, including legacy, new and mobile sensors, typically operated by different stakeholders who adopt different technological solutions. Then, raw data need to be pre-processed to guarantee data quality and mapping to common temporal and spatial references, the latter being the city graph in our case. Only after these steps can data from various sources be fused and information extracted. Finally, the information can be stored and/or be used for predictions.

We use an archetype middle-sized European city to identify and address the challenges of applying this data-to-information flow in a real-world scenario. Our archetypal city has two types of *raw traffic sensors*: (i) legacy inductive loops proving non-discriminated counts, and (ii) trajectory data from three types of mobile probes, namely taxis, buses, and individual citizens using a mobile crowdsensing application. Taxi and bus trajectory data were provided by two independent start-ups, while the mobile crowdsensing application is part of an academic project at the local university (Aguiar & Rodrigues, 2022). A challenge regarding the data sources is that probe positions were obtained at varied sampling frequencies and spatial accuracy, from various generations of Global Navigation Satellite System (GNSS) devices. *Pre-processing* tasks are necessary for each data source to (1) map-match points or trajectories onto the edges of the graph in use, providing a common spatial reference, (2) apply sanity and outlier filters, (3) interpolate to compensate for missing samples and too coarse sampling frequencies, and (4) split data into the relevant time binning, providing a common temporal reference.

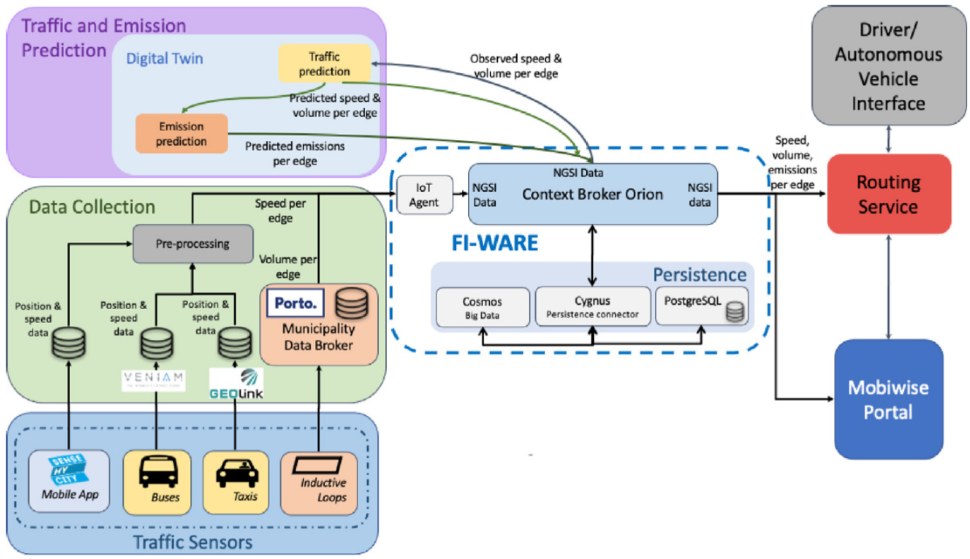

Fig. 3. Proposed microservice based architecture for data harvesting and information extraction for the MobiWise eco-routing application.

3.png 3.png

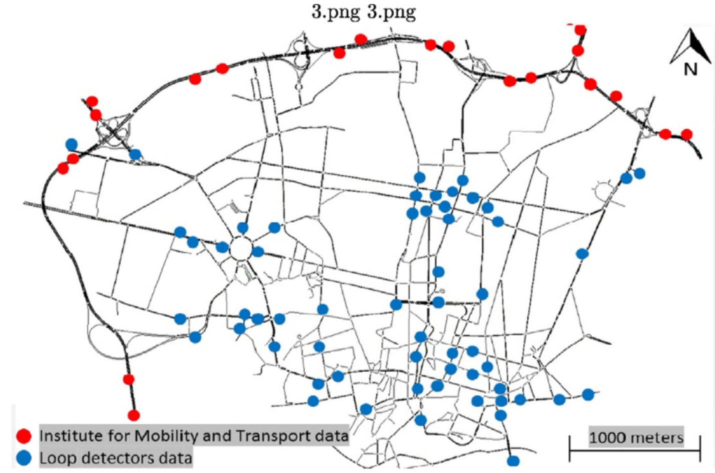

Fig. 4. Study domain with the location of data collection points (Porto, Portugal).

Finally, traffic estimations can be obtained through data fusion methods, e.g. as proposed in Silva et al. (2018), to predict congestion. In a downstream step, pollutants emissions and noise can be computed from the per-edge traffic estimations. However, these steps can only be performed for graph edges where sufficient data is available. Even after data pre-processing and fusion, traffic data often remains sparse and unevenly distributed both in space and time (Gil et al., 2017; Silva et al., 2018). This is a consequence of opportunistic sensing processes, i.e. sensing using devices and procedures whose main purpose is not to sense. An example is floating car data, as in taxis and buses, whose main purpose is to provide mobility services and not to evenly sense the city. This poses a significant challenge to data-based mobility services like traffic and road traffic emissions prediction. To address data sparsity, we use a *Digital Twin* to estimate and predict traffic on all edges of the graph used by the *Mobiwise Optimiser*. The data harvested from the various sensors are used to calibrate the *Digital Twin* instead of being directly used. Section 7 provides details on these models.

Integrating the heterogeneous data sources and producing the necessary information for the optimisation process as described above raises two main challenges: dealing with the diversity of data sources; and dealing with the different specific know-how necessary for each processing and information extraction step. The microservice architectural pattern (Cerny et al., 2018) provides the necessary level of modularity for the components, enabling independent operation by different entities and independent scalability. Standard IoT middleware technology serves as a fabric to interconnect the microservices, providing well-defined open interfaces.

A major challenge caused by data source heterogeneity is data interoperability, i.e., representing data in a common and open format (syntax), and with a common ontology to enable adequate processing (semantic). The harvested data need to be transformed into a model that provides common syntax and semantics and can be used by the computation modules, e.g. for estimating and predicting the traffic and corresponding emissions per edge. IoT middleware platforms are designed to provide the common layer in such scenarios, i.e. enabling to build large distributed systems on top of smaller heterogeneous systems in a modular and interoperable way. While ETSI is actively developing the OneM2M IoT standard[3] targeting a broad range of application areas, FI-WARE[4] is gaining traction specifically in the field of Smart Cities. We performed benchmark assessments of these two technologies with a focus on communication overhead and scalability, and FI-WARE showed better results in both cases (Aguiar & Morla,

[3] https://www.etsi.org/technologies/internet-of-things
[4] https://www.fiware.org

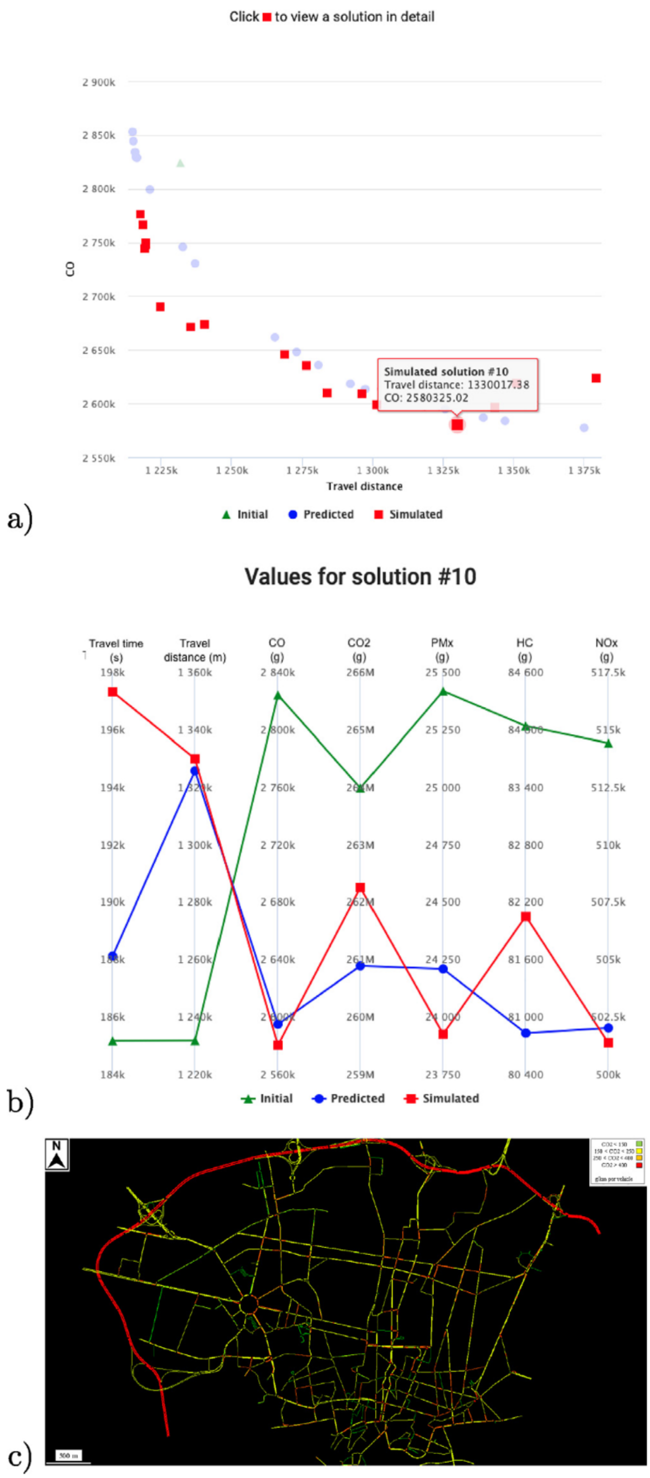

Fig. 5. Example visualisations of solutions for Scenario 1 (see Section 9 for details): (a) Pareto front with a selected solution; (b) Multiple variable graph; and (c) Heat map for CO_2 per unit distance. Note: PMx represent the sum of $PM_{2.5}$ and PM_{10}; Axis values are the summing of all eco-routing vehicles during one hour.

2019; Pereira et al., 2018). So, we next describe the interconnection of the eco-routing application components through the FI-WARE middleware, additionally identifying FI-WARE application modules that can be useful in the support of an eco-routing application. Fig. 3 represents the proposed logical architecture mapping the components identified in the context diagram of Fig. 1 to microservices for the specific case of the archetypal city, and showing how additional FI-WARE modules would blend in. The data flows among the application components are also identified.

FI-WARE addresses data interoperability through Next Generation Service Interfaces-Linked Data (NGSI-LD), also an ETSI standard (Group, 2020), as data format, and provides ontologies for smart city applications, including smart mobility. All data comes together at a data broker, FI-Ware's Orion Context Broker, which is the heart of the architecture and serves as a real-time capable interface among the microservices. The broker exposes the data through well-defined Application Programmer Interfaces (API), enabling interoperability and evolvability. Persistence modules capable of keeping historic information, FI-Ware Cygnus, can be plugged in optionally and support a large variety of storage solutions.

The publication of values for all edges of the city map represents a challenge of its own, requiring careful design and implementation of the data structures and operations used, and a scalable broker. This was thoroughly analysed in Pereira et al. (2018) and Aguiar and Morla (2019). Fi-Ware's broker Orion showed to be sufficiently agile to notify subscribed modules of a dataset corresponding to 20,000 edges, the size of our archetypal city map, when the city graph is mapped onto a database as a single data entity (Pereira et al., 2018). This article shows that how data are mapped to the naming scheme impacts the type of database operations that occurs inside the broker on each data publication and notification, which in turn is reflected in the scalability. More information about this platform is presented in Aguiar and Morla (2019).

Finally, the **Routing Service** implements the *Mobiwise Optimiser module*. This module consumes the data for the road network graph and solves the multi-objective optimisation problem (Section 5). The *Mobiwise Portal* (see Section 8) provides a visual interface for a *Decision Maker* to decide which solution best suits the desired policy. The route assignment for the chosen solution maps individual vehicle route requests, e.g. as described in Section 5, and serves them via a RESTful API to the target vehicle population. The communication among these three components can be achieved directly or via the Context Broker. However, we do not see any specific advantage in the latter since the data involved is not re-used.

7. Digital twin

Traffic data in an urban area is almost always sparse in time and space, i.e. there is no sufficiently accurate data for all edges in all time slots. Thus, we propose a digital twin to deal with the data sparsity, and describe it in this section.

7.1. Microscopic road traffic simulation

To develop an emission and noise impact, and traffic performance evaluation, the open-source microscopic traffic simulation SUMO (Simulation of Urban Mobility) version 1.7 was used (Lopez et al., 2018). This microscopic traffic simulation tool was chosen since it allows to (1) extract fully disaggregated vehicle trajectory records which can be applied to develop the estimation of pollutants and noise emissions, and traffic performance outputs; (2) model different driving behaviours parameters compliant to the road and vehicle type; (3) identify emission, noise or traffic congestion hotspots with high resolution of time and position (Lopez et al., 2018); and (4) to access and adjust driving behaviour and vehicle speed at the edge level.

The Krauss car-following model was used to characterise the vehicle driving behaviour (Treiber & Kesting, 2013; Zheng et al., 2012). It results from a modification from the Gipps car-following model, which translates into a stochastic version of the latter model. The Krauss model directly calculates the vehicle desired speed that results from a preceding determined safe speed (Treiber & Kesting, 2013; Zheng et al., 2012).

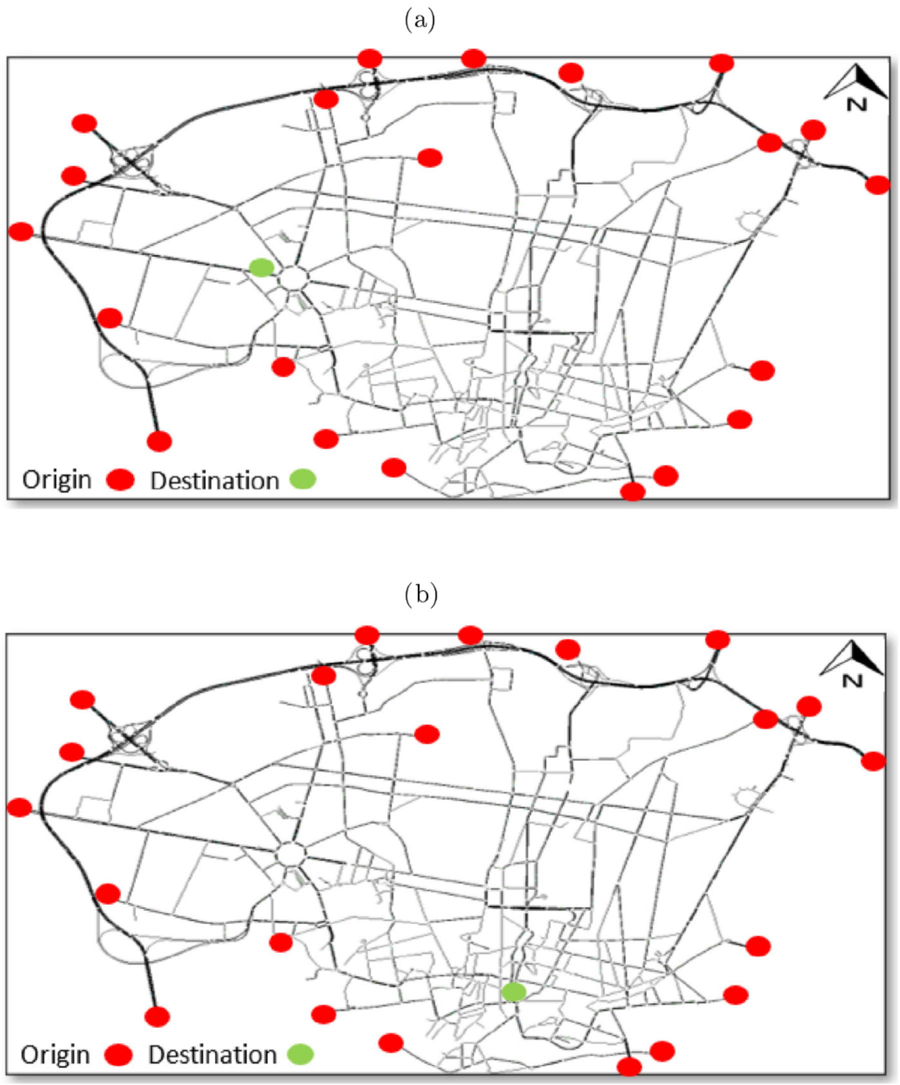

Fig. 6. Proposed eco-routing scenarios: (a) Zone 1; and (b) Zone 2.

The road simulation network in SUMO is exhibited in Fig. 4, and it comprises 2888 edges,[5] 2149 nodes, and a total length of 170 km. The Krauss car-following parameters were tuned during the calibration process to examine their effect on traffic volumes for each monitoring point (Ciuffo et al., 2012). A preliminary analysis revealed that the parameters more sensitive to driver behaviour were the minimum gap between stopped vehicles (minGap), the desired minimum time gap between the rear bumper of the leader car and the front bumper of the driver (τ), and the driver imperfection (σ) (Lopez et al., 2018).

Specific origin–destination (O-D) matrices were defined according to the period of the day and then adjusted to match traffic volumes on the monitoring points. For urban buses, their schedules, bus lanes and stop locations were defined and calibrated using the publically available information from bus companies in the metropolitan region. The widely accepted *The Geoffrey E. Havers (GEH)* statistic and regression analyses were applied to examine the consistency of the SUMO modelling platform by comparing observed and simulated traffic counts (Fernandes et al., 2019). The calibration was stopped after complying with the following calibration target — at least 85% of the monitoring points must present a GEH value below 5 for the model traffic flows (Fernandes et al., 2019).

[5] We considered only the most significant roads for this proof of concept.

7.2. Road traffic pollutant and noise emissions

The application of a COPERT-based model, as established by Macedo et al. (2020), using only the average speed (in m/s) as an input variable, allows the estimation of the pollutants emissions based on representative vehicles from the Portuguese national fleet. Such a model reduces considerably the associated computation, while still providing a good assessment of vehicular exhaust emissions for different types of vehicles (e.g., gasoline, diesel, hybrid electric) and driving conditions (urban, rural, and highway). Speed levels and emissions (in g/km) were correlated using parabolic-shaped curves for which a least-square fitting technique was used to obtain the best-fitting curves (Macedo et al., 2020).

The mentioned COPERT model was therefore used to estimate the emissions of CO_2, CO, NOx, Particulate Matter with an aerodynamic diameter of less than 2.5 micrometres ($PM_{2.5}$) and 10 micrometres (PM_{10}), and volatile organic compounds (VOCs) per edge considering the Portuguese national light-duty vehicles (LDV) fleet distribution: 39% and 40% of light duty petrol and diesel vehicles, respectively, and 21% light diesel commercial vehicles (Fernandes et al., 2019).

The Harmonoise model implemented in SUMO was used to compute the hourly equivalent continuous A-weighted sound pressure level ($L_{Aeq}{}^{1h}$) of each edge composing the network. $L_{Aeq}{}^{1h}$ represents the average energy of the fluctuating sound level on an hourly basis. More

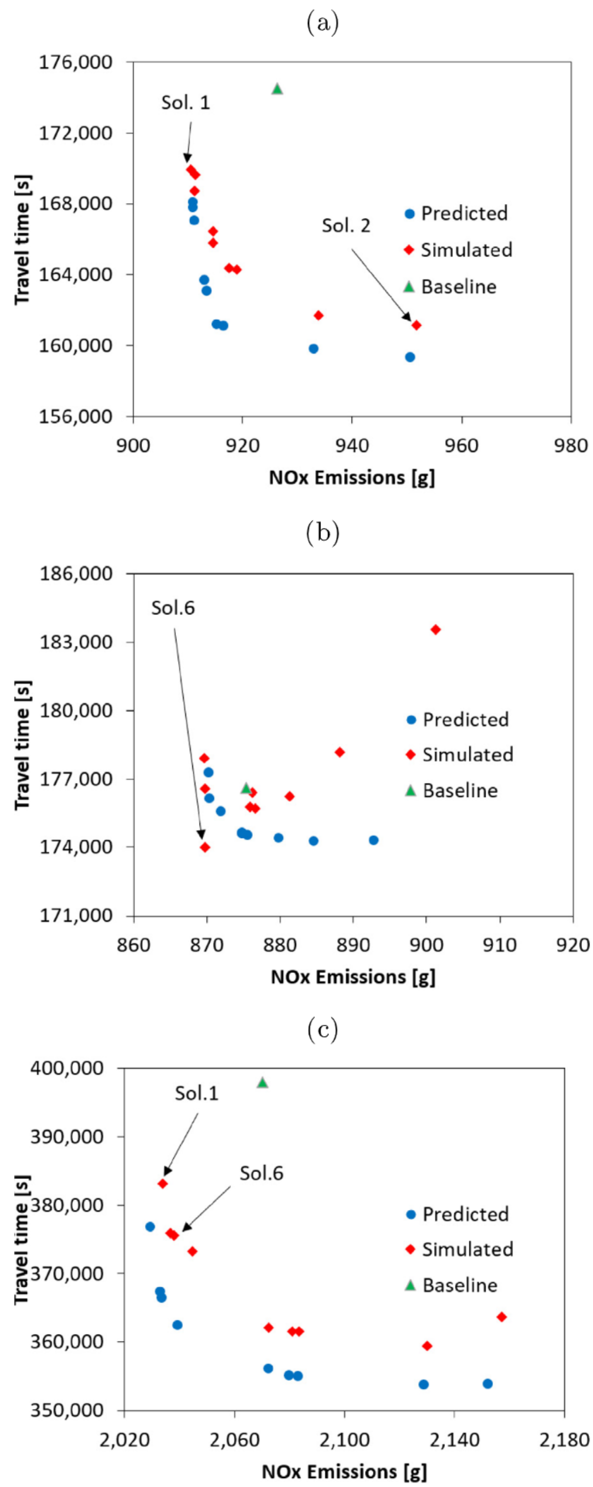

Fig. 7. The approximate Pareto fronts for NOx emissions *versus* travel time by scenario: (a) scenario 1; (b) scenario 2; and (c) scenario 3. Note: Axis values are the summing of all eco-routing vehicles during one hour along edges.

details about model formulation are given in the paper by de Vos et al. (2005).

7.3. Sustainability indicator

The application of a sustainability indicator allows to quantify the monetary cost per vehicle (EUR/veh) considering negative externalities of road traffic, such as traffic congestion, GHG and NOx emissions, noise or road crashes related costs (Korzhenevych et al., 2014). Fernandes et al. (2019) describes the methodology which provides the procedure used to achieve each component's associated contribution to the road segment aggregate cost and adjusted to local contexts of vulnerability. The calculations of each cost component by considering site-specific characteristics are provided in detail in the Appendix at the end of the article. Eq. (2) summarises external costs (EC_k) for each edge k that corresponds to the sum of each cost component (Equations 3 to 9 in Supplementary Data), as follows (Fernandes et al., 2019):

$$EC_k = TC_k + GHG_k + NOx_k + N_k + RC_k, \tag{2}$$

where EC_k is the external cost in edge k (EUR/veh); TC_k is the traffic congestion cost in edge k (EUR/veh); GHG_k is the GHG cost in edge k (EUR/veh); NOx_k is the NOx cost in edge k (EUR/veh); N_k is the noise cost in edge k (EUR/veh); RC_k is the road crash cost in edge k (EUR/veh). The estimation of RC_k can be done by applying an adjusted risk in what concerns the death and injury due to an accident for the person exposed to risk and their relatives and friends, and crash cost for the remaining society (Fernandes et al., 2019). For that purpose, crash data involving motor vehicles for the studied domain are needed.

7.4. Study area and datasets

Fig. 4 depicts the aerial view of the area of study, which serves as an archetype for the middle-sized European city. The road network belongs to the Porto Metropolitan Area, which is one of the 50 largest in the European Union, with a population of 1.74 million inhabitants (INE, 2021). An urban highway (around 9 km in length and posted speed limit of 80 km/h) connects the city downtown and its neighbourhood areas. The number of daily trips in this region was approximately 3.4 million in 2017; passenger cars and transit buses represented 68% and 11%, respectively, of these trips (INE, 2018). According to the TomTom ranking from 2019, this city exhibited average congestion levels of nearly 31%, being the second city in the national ranking where drivers spent more time in traffic (approximately 18 and 21 min per 30 min trip during morning peak and evening peak hours, respectively), as reported here (TomTom, 2022). Prior studies conducted in the region have demonstrated the negative implications of traffic emissions and noise pollution (Pascale et al., 2022; Slezakova et al., 2011), which in turn represent a matter of concern, especially for vulnerable groups. In the past few years, a smart city initiative named "Porto LivingLab" (Almeida et al., 2022; Santos et al., 2018) has developed a multi-source sensing infrastructure that captures data from several traffic, weather and environmental sensors spread in the city, making possible the testing of the proposed MobiWise eco-routing decision support system.

Road traffic flow records in 5 min-intervals were achieved through inductive loop vehicle detectors installed by the Porto City Hall. A set containing two weeks of road traffic data in April 2016 was used. Also, the average daily traffic (ADT) in urban freeways and highway roads was retrieved from the Institute for Mobility and Transport (IMT, 2020). In total, 83 monitoring points were considered.

Crash data involving motor vehicles, and motor vehicles and pedestrians were gathered from the Portuguese Road Safety Authority database for the year 2017 (ANSR, 2019). This specific year was selected since it provides the most recent dataset with precise GNSS coordinates of crash observations, which in turn is essential to assign these occurrences to a specific edge. The dataset covers a total of 927 crash observations that resulted either in injuries or fatalities.

Fig. 8. Spatial distribution of CO_2 emissions locations per unit distance in the baseline scenarios: (a) 6–7 AM.; and (b) 8–9 AM.

8. Decision support portal

The *Mobiwise Optimiser* delivers Pareto frontiers, which are a set of non-dominated solutions, i.e. solutions that cannot be improved upon without causing a reduction in another target variable. The choice of which non-dominated solution to adopt in each situation is done by the *Decision Maker*. We propose different types of visual information to support this step by illustrating the trade-offs involved in each solution, and we provide examples of those visualisations in the *Mobiwise Portal*, a website created for the purpose,[6] whose source code is publicly available.[7]

Although the optimisation problem and solver described in Section 5 can cope with any number of optimisation criteria, visualising the Pareto front in a web page raises practical challenges. Most importantly, the number of non-dominated solutions increases as the number of parameters grows. Since the portal user is a human, visualising higher dimensions and a large number of solutions presents serious challenges that fall out of the scope of this work. Thus, the current Mobwise portal version supports the choice of only two optimisation criteria because more are not supported by our visualisation proposals.

In the entrance page, the user selects optimisation parameters, among the eight criteria, CO, CO_2, NOx, PM_{10}, $PM_{2.5}$, VOC, distance and travel time.

Fig. 5 shows the visualisations available on the website. The first results displayed are the Pareto front resulting from the optimisation (predicted). Additionally, the values obtained from actually simulating the eco-routing allocation in the digital twin (simulated) are also shown in the same plot, as illustrated in Fig. 5a and carefully explained in Section 9. The plot also shows the baseline solution given by the traffic model with calibrated data and eco-routing paths based on model default (green triangle). The website user can now pick one specific solution from the Pareto front (red squares in the plot) corresponding to a specific choice of two parameters, for example, shorter travelling distance but higher CO or vice-versa.

[6] https://mobiwise.dei.uc.pt/home
[7] https://github.com/filipius/MobiWise.git

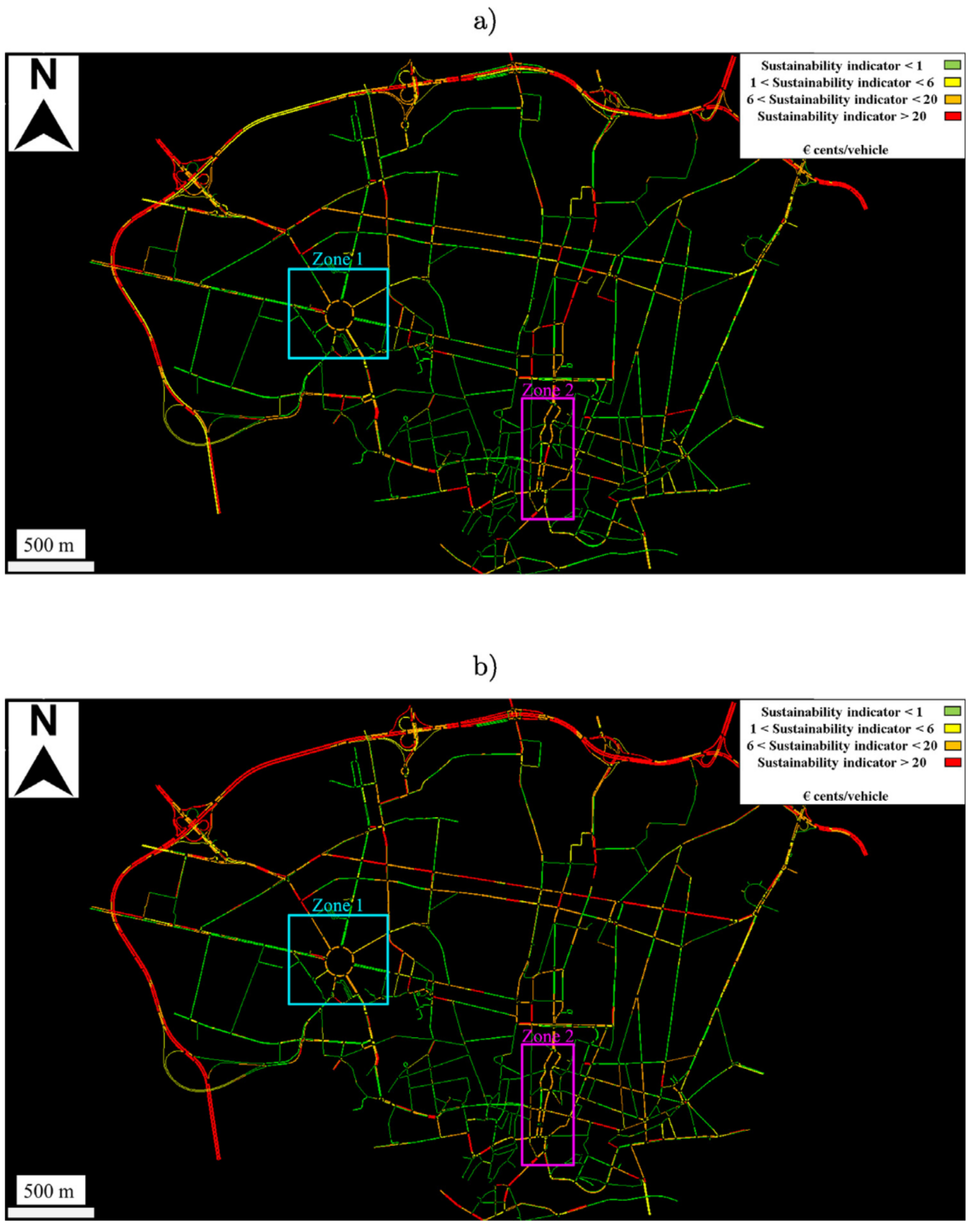

Fig. 9. Spatial distribution of external costs: (a) 6–7 AM; and (b) 8–9 AM.

A multiple variables graph, shown in Fig. 5b, appears and presents the values of all indicators. It adjusts three lines according to the particular selection: (i) the initial results without any optimisation (the green line), which never change regardless of the choice in the previous plot; (ii) the predicted results, as computed by the optimisation algorithm (the blue line); and (iii) the simulated results, obtained with the chosen solution in the digital twin (the red line). The website also reports a table with the numeric values corresponding to the multiple variables plot.

The page has two further areas of information: (i) two videos of the digital twin simulation, with the baseline simulation corresponding to the green line, and the optimised solution, corresponding to the red line; and (ii) heat maps of the city, which provide an overview of the results for different pollutants, as depicted in Fig. 5c. These heat maps enable a better understanding of the results than the videos, by illustrating with colours the graph edges that will be most affected by the optimised solution versus the baseline solution. Fig. 5c shows the distribution of CO_2 as an example.

The response time of the portal currently depends on the largest part on the time necessary to run the optimiser and generate the Pareto front (predicted), and then run the corresponding digital twin simulation (simulated). This last step allows the assessment of the impact of static link weights, which is relevant mostly for research purposes. The delay for generating the Pareto front in the Optimiser is around 2 s.[8] In a real system, only these solutions will be calculated, and not the simulated ones. Moreover, we believe that the decision maker will find this response time reasonable.

9. Calibration and results

9.1. Eco-routing scenarios

The main idea behind the eco-routing scenarios is to simulate the existence of centralised traffic management advising a sub-population of eco-routing commuters using calibrated traffic data from our archetypal Porto city's road network. The proposed scenarios focus on the analysis

[8] Values obtained on Intel Core i5 with 6 MB cache and 8 GB RAM.

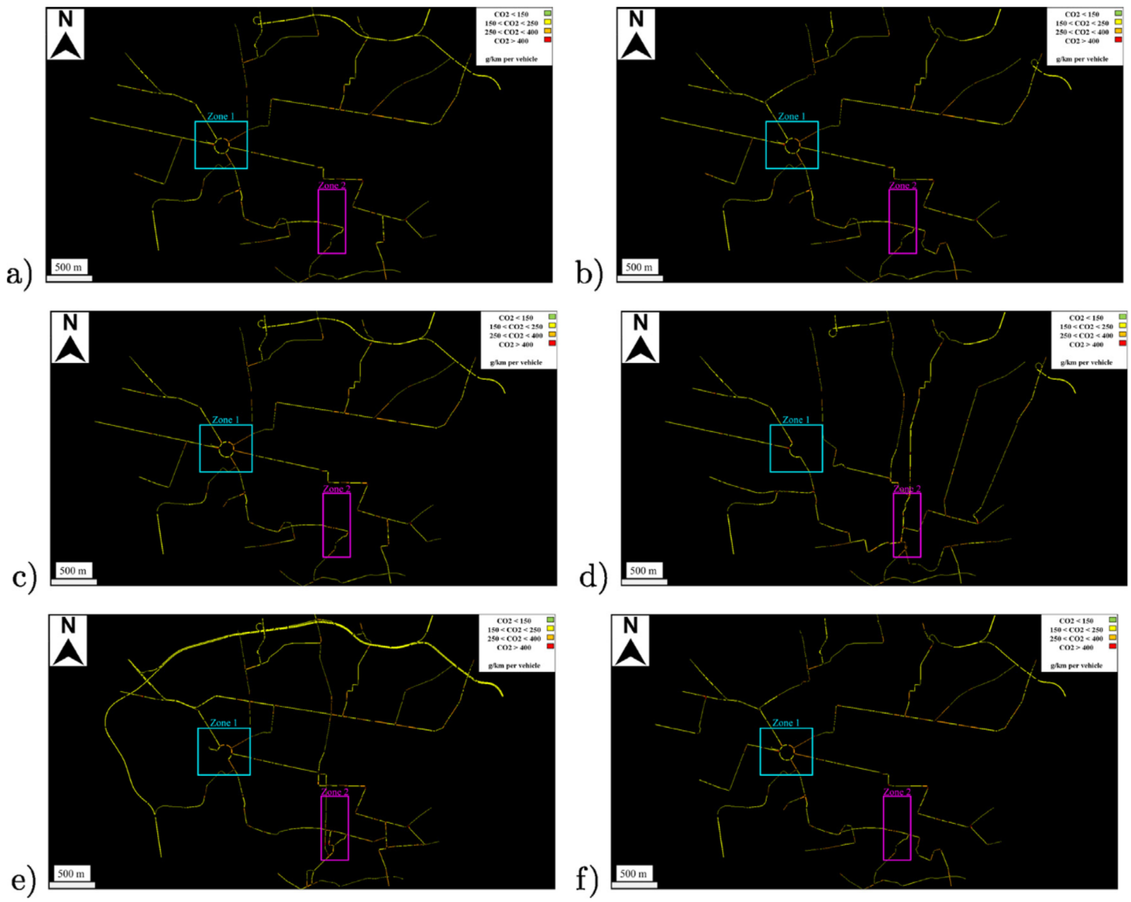

Fig. 10. Spatial distribution of CO_2 emissions locations per unit distance by scenario: (a) baseline routing solution — Zone 1 at 6–7 AM; (b) optimal routing solution — Zone 1 at 6–7 AM; (c) baseline routing solution — Zone 2 at 6–7 AM; (d) optimal routing solution — Zone 2 at 6–7 AM; (e) baseline routing solution — Zone 1 at 8–9 AM; and (f) optimal routing solution — Zone 2 at 8–9 AM.

of vehicles coming from multiple origins to the same destination, the latter one being a city hub, such as a train station, or commercial and business areas. A fixed percentage of vehicles entering the simulation of the network from centroids was redistributed along new routes based on the eco-routing optimisation results. The following scenarios were examined:

- **Scenario 1** - 5% of eco-routing commuters from centroids to a city commercial and cultural areas (Zone 1, Boavista in the Mobiwise portal) between 6–7 AM (Fig. 6a);
- **Scenario 2** - 5% of eco-routing commuters from centroids to the city train station (Zone 2, São Bento in the Mobiwise portal) between 6–7 AM (Fig. 6b);
- **Scenario 3** - Similar to scenario 1, but the simulation experiments were done during the peak hour (8–9 AM period).

The above-mentioned scenarios were compared with a baseline scenario that represents the traffic conditions in the study area with a validated model at 6–7 AM and 8–9 AM, where all vehicles have assigned paths to the specific destination. For these cases, the simulation comprises vehicles with the usual origin–destination matrix, considering 7,300 and 20,000 vehicles at 6–7 AM and 8–9 AM, respectively. The analysis of the potential benefits of the eco-routing scenarios was performed for both the overall network and on an edge basis.

9.2. SUMO calibration and validation

This section presents the main results concerning the traffic model calibration and validation. Both procedures were done for 1 h because information retrieved from the Institute for Mobility and Transport (IMT) is given daily. To obtain hourly counts, loop sensor data in the study region were assigned in 24 intervals of 1 h to further get the relative contribution (in percentage) of each period on the total daily traffic. Finally, these values were used to compute hourly traffic from the IMT monitoring points. The following calibrated SUMO model parameters were obtained: minGap = 1.20 m; τ = 1.20 s; and σ = 0.90. The validation of traffic volumes in 83 evaluation points along the study area and by demand period showed as robust. The coefficient of determination values (R^2) between estimated (SUMO traffic) and observed data were higher than 0.90, regardless of the studied period. Approximately 94% and 90% of the loop detectors at 6–7 AM and 8–9 AM, respectively, achieved GEH values lower than 4, meaning that the simulated and the observed traffic data for the eco-routing case is accurate (Fernandes et al., 2019). These calibrated settings were subsequently applied in the eco-routing scenarios.

9.3. Optimisation

For validation purposes, each Pareto-optimal solution was simulated in the digital twin, and the costs obtained in each simulation were compared to the predicted costs given by the objective values of the minimum-cost flow problem for each solution. This comparison allows assessing the robustness of the model that is based on average values to predict the cost solution, while the digital twin performs a microscopic simulation for that solution. For that purpose, each of the Pareto-optimal flows was converted into a set of routes as described at the end of Section 5 which were then evaluated using the digital twin.

Figs. 7a–c exhibit the initial baseline scenario, the optimal Pareto frontier, and the results for a SUMO simulation of each solution on the

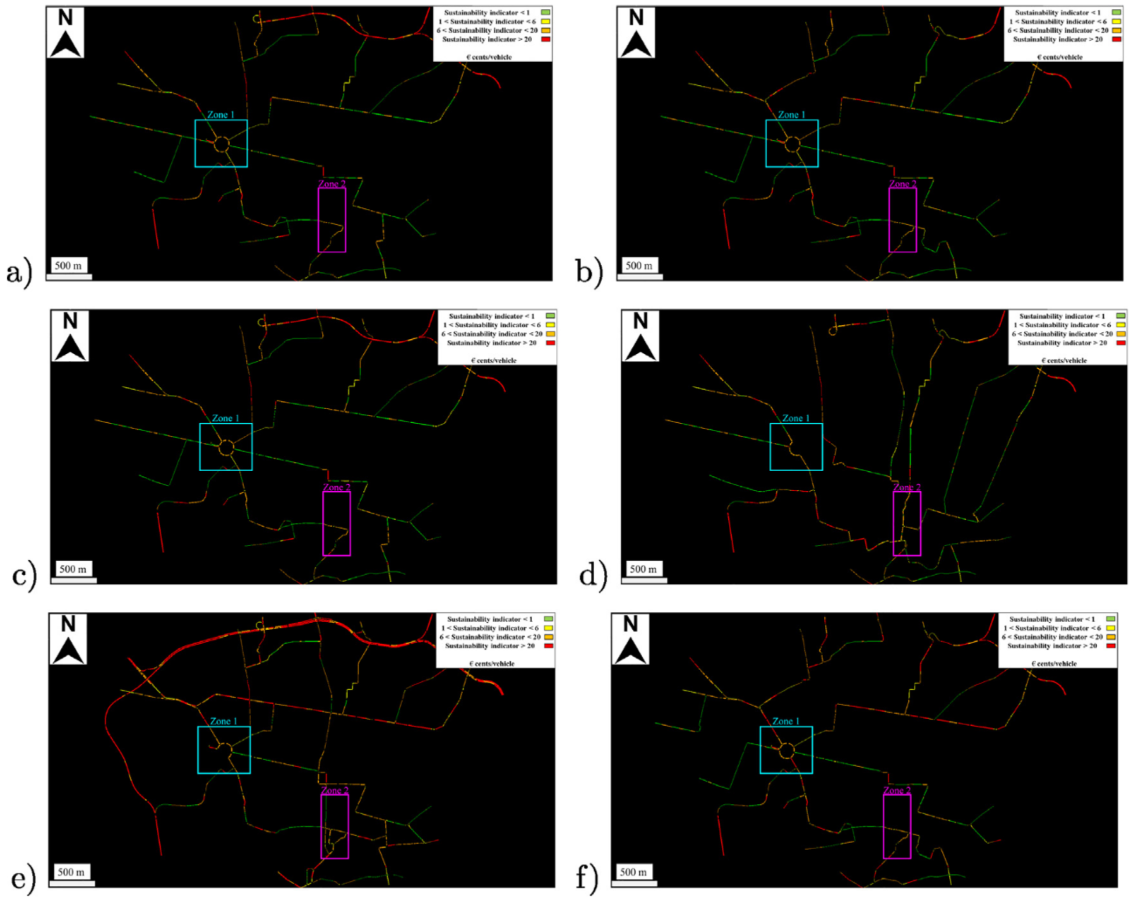

Fig. 11. Spatial distribution of the external costs by scenario: (a) baseline routing solution — Zone 1 at 6–7 AM; (b) optimal routing solution — Zone 1 at 6–7 AM; (c) baseline routing solution — Zone 2 at 6–7 AM; (d) optimal routing solution — Zone 2 at 6–7 AM; (e) baseline routing solution — Zone 1 at 8–9 AM; and (f) optimal routing solution — Zone 1 at 8–9 AM.

Pareto frontier by the demand scenario. The selected optimised criteria were travelled time *versus* NOx emissions of eco-routing vehicles. Other optimised criteria were also tested, such as travel time *versus* distance or travel time *versus* CO_2, as listed in Table 1. However, the resulting gains were not so notable as NOx *versus* travel time criteria did. The difference between the optimal solution (predicted) and the one obtained via simulation on the digital twin is due to the use of static weights on the edges in the optimal solutions, i.e. during the optimisation process the changes caused by the re-distribution of the vehicles subject to eco-routing on each edge are not considered.

For scenario 1 (Fig. 7a), the findings confirmed that all solutions reduced eco-routing travel times relative to the baseline solution; however, this outcome did not hold for NOx. For solution 2 (which is closest to the abscissa of the graph) travel time decreased by 7.7% but NOx emissions could increase up to 2.8%. Such a solution is characterised by the diversion of traffic to faster but longer routes. Although predicted and simulated Pareto fronts yielded different travel time values, the differences between NOx optimal values were small.

Results for scenario 2 revealed that several optimal solutions with the lowest emissions increased the travel time in relation to the existing conditions, as exhibited in Fig. 7b. This scenario recorded different values between predicted and simulated optimal data for both criteria, which is mostly explained by the limited number of paths to the Zone 2 destination (São Bento). This is explained in detail in Section 9.4.

Higher improvements can be also seen in scenario 3, which is associated with peak hour conditions (Fig. 7c). For instance, if decision-makers adopted solution 6, then eco-routing vehicles could save up to 1.6% and 5.6% in their NOx and travel time, respectively. Pareto fronts also showed that solutions aiming at minimising eco-routing travel time resulted in additional increases of NOx emissions (up to 4%) in scenario 3 compared to the baseline traffic conditions.

Table 1
Number of Pareto optimal solutions.

Scenario	Optimised criteria		
	Time *vs* Distance	Time *vs* NO_x	Time *vs* CO_2
1	12	9	7
2	11	9	8
3	14	9	9

9.4. Emissions and sustainability indicator results

This section first presents the results regarding the baseline scenarios of Zone 1 between off-peak and peak hours, followed by a comparative analysis between baseline and optimal solutions in all testing scenarios.

Fig. 8a–b shows the hot-spot CO_2 emissions locations per unit distance in the baseline scenarios of Porto Zone 1 6–7 AM and 8–9 AM periods. The urban freeway across the study domain accounts for a great part of the pollutant emissions in the network, i.e., approximately 36% and 46% of the CO_2 and NO_x network emissions for the 6–7 AM and 8–9 AM periods, respectively. This road section only represents 10% of the network edges' total distance. It was also observed that, in the main network, signalised intersections and roundabouts achieved the highest values of CO_2 emissions locations per unit distance with an average value 20% higher than the network average CO_2 value.

The average edge travel time by vehicle for baseline scenarios was about 6 and 7 s for the 6–7 AM period and 8–9 AM, respectively. The percentage of edges with travel time higher than 8 and 12 s was 23%–25% (depending on the scenario) and 14%–15% (depending on the scenario), respectively. A detailed analysis of the travel time confirmed that the edges with the highest travel time were located near traffic lights. Vehicles spent time at mid-block areas between closely-spaced intersections (spacing lower than 200 m) mostly due to the high traffic volumes that did not allow a good progression through the intersections.

Road crash (RC) and traffic congestion (TC) costs are presented as the largest contributors to the sustainability indicator network-wide by accounting for more than 95% of the share. Edges with the highest external costs were located on the urban freeway (see Fig. 9a–b). As suspected, the number of edges with red colour, i.e., EC higher than 20 ct/veh is higher during the 8–9 AM period. This outcome is possibly explained by the higher contribution of TC; its contribution to external costs rose from 10% to 17% between 6–7 AM and 8–9 AM periods.

Figs. 10 and 11 depict CO_2 emissions per unit distance and external costs (i.e., sustainability indicator) hot-spots on the edges related to the routes completed by the routing vehicles for both baseline and eco-routing scenarios. For the analysis, we selected Pareto Front solutions number 1, 6 and 1 of Scenarios 1, 2 and 3, respectively (see Fig. 7 for solutions visualisation). For the off-peak period (6–7 AM), the optimisation of routing vehicles led to a decrease in the amount of pollutant emissions from these vehicles, e.g., CO_2 and NO_x by 1.8% and 0.7% at Zone 1 and Zone 2, respectively. It must be noted that Zone 2 is located in the downtown area of the city where few paths are available for most of the eco-routing vehicles to reach this destination, which explains these lower gains. However, all indicators were improved after diverting eco-routing vehicles to alternative routes. The optimal solution also showed to be effective in reducing eco-routing vehicles specific emissions at peak hours (8–9 AM); both CO_2 and NO_x decreased by 2.1% and 1.8%, respectively, in relation to the baseline conditions. Although solutions were based on NOx and travel time criteria, reductions up to 1.2% in the travelled distance can be expected.

Concerning the overall network, it can be noted that Scenarios 1 and 2 decreased emissions and travel time in less than 0.2%, while Scenario 3 achieved reductions of nearly 0.5%, 0.4% and 0.7% in CO_2 and NO_x and travel time, respectively. From the results reported in Scenario 3, assuming just 1 h per day along a year for the same travel demand, it is possible to save approximately 400 tons of CO_2 and 1.2 tons of NO_x, which seems significant in the study area.

Since the routing optimisation of these vehicles did not consider all components of the sustainability indicator, these solutions provided routes exhibiting higher external costs (~3%). Fig. 11 displays several edges with orange (EC > 6 ct/veh) or red (EC > 20 ct/veh) colours in the optimal solutions. This happened because the traffic was diverted to other routes that are composed of edges with a moderate frequency of road crashes, leading to an increase in RC.

Despite the small benefits, the potential of the application of our tools is demonstrated, since emission impacts travel time, and travelled distance was simultaneously reduced. It is worth noticing that digital twin benefits were obtained by only considering 5% of traffic diverted to other routes, which is relevant. If the routes related to a higher percentage of vehicles were optimised by our tools, then significant gains would be expected. Accordingly, traffic planners can use this methodology to choose the most suitable criterion (or criteria) according to the network-specific needs.

10. Conclusions

Smart cities have the potential to drastically reduce road traffic-related externalities, i.e., climate change (CO_2), air pollution, traffic congestion, road crashes and noise, and in the end, improve the population's quality of life. The emergence of the IoT provides technological tools that enable harvesting and harmonising the necessary data, as well as the integration of multiple components into complex smart mobility applications. This work provides the blueprint of one such application — eco-routing for multiple origin single destination scenarios. We applied a use-case-driven design methodology to derive the system components and describe each component and the potential of IoT as enabling technology. We also propose the use of a traffic micro-simulator calibrated with real data as a digital twin to deal with traffic data sparsity. Our results obtained with the digital twin show that route optimisation for 5% of the vehicles can reduce total pollutant emissions up to 2.1% in 1 h in a mid-sized European city. This re-distribution of traffic during 1 h per day would correspond to savings of approximately 400 tons of carbon dioxide and 1.2 tons of nitrogen oxides throughout a year. We also show that achievable reductions depend on the existence of alternative routes for traffic distribution.

One scientific contribution of the study is that it introduces a flexible eco-friendly and sustainable routing service based on IoT technologies that cover social, environmental, and economic sustainable factors by considering the site-specific needs in an integrated way. These characteristics include but are not limited to road type, traffic control treatment, traffic signal plans, speed limits, vehicle types, car fleet distributions, population density, congestion levels, or several circulating lanes. Such a system is flexible enough to incorporate other criteria parameters or to optimise transport system operations for certain transport modes to minimise traffic congestion, pollution and road safety impacts at both system and road levels. From a societal perspective, this research can endow vehicle navigation systems with flexible analysis of road traffic-related impacts that can accommodate city-specific needs for drivers, population and travellers. We can incorporate the proposed methodology into a traffic management tool that helps decision-makers to identify certain polluting, noisy and black crash hotspot locations in cities. This can support future traffic restricting measures and taxing road strategies, for example, defining cost thresholds for a given externality during short periods of high traffic demand, such as events, road accidents, and access to a busy shopping mall.

This work has several limitations: traffic model calibration was done manually; optimisation was run based on historic data, although the proposed architecture enables online adaptation; eco-routing management system was tested for one real-world case study, which limits the generalisation of study findings for regions with identical size, travelling distances, available public and private modes, presence of heterogeneous and legacy sensors, vehicle fleet composition or demand levels; emissions were estimated using a model based on average speed, which will not identify emission peaks at the entrance and exit of intersections and highways; impacts of PMx on human health were not considered in the eco-indicator calculation. From a research perspective, this paper opens several paths for future work: (1) evolution of multi-objective optimisation to a model with dynamic weights, along with the development of closed-form models for the various pollutant emissions that could be embedded in the optimisation step; (2) automated calibration of traffic micro-simulation digital twins with data collected from the city sensing infrastructure in real time; (3) addition of the impact on air quality and related impacts; (4) addition of soft modes (pedestrians, cyclists and motorcyclists) to the digital twin and eco-routing alternatives; and (5) testing the proposed eco-routing system in other digital twins with variations in network size, traffic volumes, fleet compositions, directional distributions, and percentage of eco-routing vehicles.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Portal code open source, and plots visualisable there for presented scenarios. Other data cannot be shared.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.scs.2022.104180.

References

Aguiar, A., & Morla, R. (2019). Lessons learned and challenges on benchmarking publish-subscribe IoT platforms. In *Proceedings of the 2nd workshop on benchmarking cyber-physical systems and internet of things - CPS-IoTBench '19* (pp. 24–29). New York, New York, USA: ACM Press.

Aguiar, A., & Rodrigues, J. (2022). SenseMyCity: A mobile IoT tool for researching intelligent urban mobility. In *2022 14th international conference on communication systems & networks* (pp. 725–733). IEEE.

Ahn, K., & Rakha, H. (2008). The effects of route choice decisions on vehicle energy consumption and emissions. *Transportation Research Part D: Transport and Environment*, *13*(3), 151–167.

Ahuja, R. K., Magnanti, T. L., Orlin, J. B., & Weihe, K. (1995). Network flows: Theory, algorithms and applications. *ZOR-Methods and Models of Operations Research*, *41*(3), 252–254.

Almeida, A., Brás, S., Oliveira, I., & Sargento, S. (2022). Vehicular traffic flow prediction using deployed traffic counters in a city. *Future Generation Computer Systems*, *128*, 429–442.

ANSR (2019). *Sinistralidade rodoviária - annual report (in Portuguese)*: *Technical report*, Portuguese Road Safety Authority.

Arsenio, E., Silva, S., Iglésias, H. P., & Domingues, A. (2020). Assessing multimodal mobility trends using heterogeneous data sources: A case study for supporting sustainable policy goals within the region of algarve. *Transportation Research Procedia*, *49*, 107–118.

Bandeira, J. M., Guarnaccia, C., Fernandes, P., & Coelho, M. C. (2018). Advanced impact integration platform for cooperative road use. *International Journal of Intelligent Transportation Systems Research*, *16*(1), 1–15.

Bandeira, J. M., Tafidis, P., Macedo, E., Teixeira, J., Bahmankhah, B., Guarnaccia, C., & Coelho, M. C. (2020). Exploring the potential of web based information of business popularity for supporting sustainable traffic management. *Transport and Telecommunication Journal*, *21*(1), 47–60.

Bansal, M., Chana, I., & Clarke, S. (2020). A survey on IoT big data: Current status, 13 V's challenges, and future directions. *ACM Computing Surveys*, *53*(6).

Boriboonsomsin, K., Barth, M. J., Zhu, W., & Vu, A. (2012). Eco-routing navigation system based on multisource historical and real-time traffic information. *IEEE Transactions on Intelligent Transportation Systems*, *13*(4), 1694–1704.

Boriboonsomsin, K., Dean, J., & Barth, M. (2014). Examination of attributes and value of ecologically friendly route choices. *Transportation Research Record: Journal of the Transportation Research Board*, *2427*(1), 13–25.

Calderón, F., & Miller, E. J. (2020). A literature review of mobility services: definitions, modelling state-of-the-art, and key considerations for a conceptual modelling framework. *Transport Reviews*, *40*(3), 312–332.

Cerny, T., Donahoo, M. J., & Trnka, M. (2018). Contextual understanding of microservice architecture: Current and future directions. *SIGAPP Appllied Computer Reviews*, *17*(4), 29–45.

Ciuffo, B., Punzo, V., & Marcello, M. (2012). *The calibration of traffic simulation models : Report on the assessment of different Goodness of Fit measures and Optimization Algorithms MULTITUDE Project – COST Action TU0903*: *Technical report*, European Union.

Cockburn, A. (1997). Structuring use cases with goals. *Journal of Object-Oriented Programming*, *10*(5), 56–62.

de Vos, P., Beuving, M., & Verheije, E. (2005). Harmonised accurate and reliable methods for the EU directive on the assessment and management of environmental noise. *Final Technical Report*, AEA Technology Rail BV.

Djavadian, S., Tu, R., Farooq, B., & Hatzopoulou, M. (2020). Multi-objective eco-routing for dynamic control of connected & automated vehicles. *Transportation Research Part D: Transport and Environment*, *87*, Page 102513.

Elbery, A., & Rakha, H. (2019). City-wide eco-routing navigation considering vehicular communication impacts. *Sensors*, *19*(2), 290.

Ericsson, E., Larsson, H., & Brundell-Freij, K. (2006). Optimizing route choice for lowest fuel consumption – Potential effects of a new driver support tool. *Transportation Research Part C (Emerging Technologies)*, *14*(6), 369–383.

European Environment Agency (2020). *Air quality in Europe - 2020 report*: *Technical report*, European Environment Agency.

Fernandes, P., Teixeira, J., Guarnaccia, C., Bandeira, J. M., Macedo, E., & Coelho, M. C. (2018). The potential of metering roundabouts: Influence in transportation externalities. *Transportation Research Record*, *2672*(25), 21–34.

Fernandes, P., Vilaça, M., Macedo, E., Sampaio, C., Bahmankhah, B., Bandeira, J., Guarnaccia, C., Rafael, S., Fernandes, A., Relvas, H., Borrego, C., & Coelho, M. (2019). Integrating road traffic externalities through a sustainability indicator. *Science of the Total Environment*, *691*, 483–498.

Firdhous, M. F. M., Sudantha, B. H., & Hussien, N. A. (2021). A framework for IoT-enabled environment aware traffic management. *International Journal of Electrical and Computer Engineering (IJECE)*, *11*(1), 518–527.

Gil, D. S., D'Orey, P. M., & Aguiar, A. (2017). On the challenges of mobile crowdsensing for traffic estimation. In *Proceedings of the 15th ACM conference on embedded network sensor systems* (pp. 1–2). New York, NY, USA: ACM.

Group, C. I. M. E. I. S. (2020). ETSI GS CIM 009: Context information management (CIM); NGSI-LD API.

Guido, G., Rogano, D., Vitale, A., Astarita, V., & Festa, D. (2017). Big data for public transportation: A DSS framework. In *2017 5th IEEE international conference on models and technologies for intelligent transportation systems* (pp. 872–877). IEEE.

Hamacher, H. W., Pedersen, C. R., & Ruzika, S. (2007). Multiple objective minimum cost flow problems: A review. *European Journal of Operational Research*, *176*(3), 1404–1422.

Hussain, D., Khan, M. A., Abbas, S., Naqvi, R. A., Mushtaq, M. F., Rehman, A., & Nadeem, A. (2021). Enabling smart cities with cognition based intelligent route decision in vehicles empowered with deep extreme learning machine. *Computers, Materials & Continua*, *66*(1), 141–156.

IMT (2020). *Relatório de tráfego na rede nacional de auto-estradas (In Portuguese)*: *Technical report*, Portuguese Institute for Mobility and Transport.

INE (2018). *Mobilidade e funcionalidade do território nas Áreas metropolitanas do porto e de lisboa (In Portuguese)*: *Technical report*, Statistics of Portugal.

INE (2021). *Variação da população residente 2011–2021, Munícipios - Continente (In Portuguese)*: *Technical report*, Statistics of Portugal.

Korzhenevych, A., Dehnen, N., Bröcker, J., Holtkamp, M., Meier, H., Gibson, G., Varma, A., & Cox, V. (2014). *Update of the handbook on external costs of transport - final report*: *Technical report*, European Commission - DG Mobility and Transport.

Li, B., Xu, S., & Peng, H. (2020). Eco-routing for plug-in hybrid electric vehicles. In *2020 IEEE 23rd international conference on intelligent transportation systems* (pp. 1–6).

Lopez, P. A., Behrisch, M., Bieker-Walz, L., Erdmann, J., Flötteröd, Y.-P., Hilbrich, R., Lücken, L., Rummel, J., Wagner, P., & Wießner, E. (2018). Microscopic traffic simulation using SUMO. In *IEEE international conference on intelligent transportation systems* (pp. 2575–2582). IEEE.

Ma, J., Hu, J., Leslie, E., Zhou, F., Huang, P., & Bared, J. (2019). An eco-drive experiment on rolling terrains for fuel consumption optimization with connected automated vehicles. *Transportation Research Part C (Emerging Technologies)*, *100*, 125–141.

Macedo, E., Tomás, R., Fernandes, P., Coelho, M., & Bandeira, J. (2020). Quantifying road traffic emissions embedded in a multi-objective traffic assignment model. In *Transportation research procedia, vol. 47* (pp. 648–655).

Majumdar, S., Subhani, M. M., Roullier, B., Anjum, A., & Zhu, R. (2021). Congestion prediction for smart sustainable cities using IoT and machine learning approaches. *Sustainable Cities and Society*, *64*, Page 102500.

Melkonyan, A., Gruchmann, T., Lohmar, F., & Bleischwitz, R. (2022). Decision support for sustainable urban mobility: A case study of the rhine-ruhr area. *Sustainable Cities and Society*, *80*, Page 103806.

Mondal, M. A., & Rehena, Z. (2021). An IoT-based congestion control framework for intelligent traffic management system. In N. N. Chiplunkar, & T. Fukao (Eds.), *Advances in artificial intelligence and data engineering* (pp. 1287–1297). Singapore: Springer Singapore.

Orfila, O., Saint Pierre, G., & Messias, M. (2015). An android based ecodriving assistance system to improve safety and efficiency of internal combustion engine passenger cars. *Transportation Research Part C (Emerging Technologies)*, *58*, 772–782.

Pascale, A., Mancini, S., d'Orey, P. M., Guarnaccia, C., & Coelho, M. C. (2022). Correlating the effect of COVID-19 lockdown with mobility impacts: A time series study using noise sensors data. *Transportation Research Procedia*, *62*, 115–122.

Pereira, C., Cardoso, J., Aguiar, A., & Morla, R. (2018). Benchmarking Pub/Sub IoT middleware platforms for smart services. *Journal of Reliable Intelligent Environments*, *4*(1), 25–37.

Pirra, M., & Diana, M. (2019). Integrating mobility data sources to define and quantify a vehicle-level congestion indicator: An application for the city of Turin. *European Transport Research Review*, *41*(11).

Priya, S. C., Kumar, S., & Prabhakaran, S. (2022). Eco-routing—to reduce vehicle CO2 emissions by CACC: An IoT application. In S. Chandramohan, B. Venkatesh, S. Sekhar Dash, S. Das, & C. Sharmeela (Eds.), *Artificial intelligence and evolutionary computations in engineering systems* (pp. 81–93). Singapore: Springer Singapore.

Rhun, A. L., Bonnans, F., Nunzio, G. D., Leroy, T., & Martinon, P. (2020). An eco-routing algorithm for HEVs under traffic conditions. *IFAC-PapersOnLine*, *53*(2), 14242–14247, 21st IFAC World Congress.

Santos, P. M., Rodrigues, J. G. P., Cruz, S. B., Lourenço, T., d'Orey, P. M., Luis, Y., Rocha, C., Sousa, S., Crisóstomo, S., Queirós, C., Sargento, S., Aguiar, A., & Barros, J. (2018). PortoLivingLab: An IoT-based sensing platform for smart cities. *IEEE Internet of Things Journal*, *5*(2), 523–532.

Shen, L., Shao, H., Wu, T., Lam, W. H., & Zhu, E. C. (2019). An energy-efficient reliable path finding algorithm for stochastic road networks with electric vehicles. *Transportation Research Part C (Emerging Technologies)*, *102*, 450–473.

Silva, C., d'Orey, P. M., & Aguiar, A. (2018). Interpreting traffic congestion using fundamental diagrams and probabilistic graphical modeling. In *2018 IEEE international conference on data mining workshops* (pp. 580–587). IEEE.

Silva, C., Orey, P. M., & Aguiar, A. (2019). Visual analysis of multivariate urban traffic data resorting to local principal curves. In D. Archambault, I. Nabney, & J. Peltonen (Eds.), *Machine learning methods in visualisation for big data*. The Eurographics Association.

Sivak, M., & Schoettle, B. (2012). Eco-driving: Strategic, tactical, and operational decisions of the driver that influence vehicle fuel economy. *Transport Policy*, *22*, 96–99.

Slezakova, K., Castro, D., Begonha, A., Delerue-Matos, C., da Conceição Alvim-Ferraz, M., Morais, S., & do Carmo Pereira, M. (2011). Air pollution from traffic emissions in oporto, Portugal: Health and environmental implications. *Microchemical Journal*, *99*(1), 51–59.

Sun, J., & Liu, H. X. (2015). Stochastic eco-routing in a signalized traffic network. *Transportation Research Procedia*, *7*, 110–128.

Texas A& M. Transportation Institute (2021). *2021 urban mobility report*: *Technical report*, The Texas A&M Transportation Institute.

Tomás, R., Fernandes, P., Macedo, J., & Coelho, M. C. (2021). Carpooling as an immediate strategy to post-lockdown mobility: A case study in university campuses. *Sustainability*, *13*(10), Page 5512.

Treiber, M., & Kesting, A. (2013). Car-following models based on driving strategies. In *Traffic flow dynamics: Data, models and simulation* (pp. 181–204). Berlin, Heidelberg: Springer Berlin Heidelberg.

Wang, J., Elbery, A., & Rakha, H. A. (2019a). A real-time vehicle-specific eco-routing model for on-board navigation applications capturing transient vehicle behavior. *Transportation Research Part C (Emerging Technologies)*, *104*, 1–21.

You, L., Zhao, F., Cheah, L., Jeong, K., Zegras, P. C., & Ben-Akiva, M. (2020). A generic future mobility sensing system for travel data collection, management, fusion, and visualization. *IEEE Transactions on Intelligent Transportation Systems*, *21*(10), 4149–4160.

Zeng, W., Miwa, T., & Morikawa, T. (2020). Eco-routing problem considering fuel consumption and probabilistic travel time budget. *Transportation Research Part D: Transport and Environment*, *78*, Page 102219.

Zhang, J., & He, S. (2020). Smart technologies and urban life: A behavioral and social perspective. *Sustainable Cities and Society*, *63*, Page 102460.

Zhang, Q., Jian, D., Xu, R., Dai, W., & Liu, Y. (2017). Integrating heterogeneous data sources for traffic flow prediction through extreme learning machine. In *2017 IEEE international conference on big data* (pp. 4189–4194).

Zhao, X., Wu, Y., Rong, J., & Zhang, Y. (2015). Development of a driving simulator based eco-driving support system. *Transportation Research Part C (Emerging Technologies)*, *58*, 631–641.

Zheng, J., Suzuki, K., & Fujita, M. (2012). Evaluation of car-following models using trajectory data from real traffic. *Procedia - Social and Behavioral Sciences*, *43*, 356–366.

Zhou, M., Jin, H., & Wang, W. (2016). A review of vehicle fuel consumption models to evaluate eco-driving and eco-routing. *Transportation Research Part D: Transport and Environment*, *49*, 203–218.