FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

Data-Driven Pedestrian Mobility Assessment for Urban Monitoring Platforms

Maria Pinto Calçada

Mestrado em Engenharia Eletrotécnica e de Computadores

FEUP Supervisor: Prof. Daniel G. Costa

July 14, 2026

Data-Driven Pedestrian Mobility Assessment for Urban Monitoring Platforms

Maria Pinto Calçada

Mestrado em Engenharia Eletrotécnica e de Computadores

Approved in oral examination by the committee:

President: Prof. Maria Teresa Magalhães da Silva Pinto de Andrade
Examiner: Prof. Sara Teixeira Baltazar
Member: Prof. Daniel Gouveia Costa

July 14, 2026

Resumo

As cidades recorrem cada vez mais a plataformas digitais para monitorizar o tráfego, os transportes públicos, a qualidade do ar, entre outros sistemas urbanos. No entanto, um dos aspetos mais comuns da mobilidade diária permanece frequentemente menos visível: a experiência dos peões. Em particular, o tempo de espera em passadeiras com semáforos pode interromper os percursos a pé, reduzir a continuidade de deslocações curtas e tornar a circulação pedonal menos confortável. Embora estes atrasos ocorram em pontos específicos da cidade, raramente são traduzidos em camadas espaciais que possam ser inspecionadas e atualizadas pelas equipas de gestão urbana.

Esta dissertação centra-se nos cruzamentos urbanos semaforizados do Porto, onde o tempo de espera dos peões pode afetar diretamente as condições de circulação a pé. Para estudar este problema, foi desenvolvida uma metodologia que combina dados espaciais, atrasos medidos dos peões, simulação de percursos , e classificação baseada nos atrasos. A abordagem proposta combina dados espaciais abertos, valores de atraso medidos no terreno, simulação de percursos pedestres locais e classificação Level of Service (LOS) baseada no atraso. Os dados OpenStreetMap (OSM) são utilizados para extrair a rede rodoviária, identificar cruzamentos, detetar passagens para peões e distinguir cruzamentos com controlo semafórico para peões. As observações no terreno são então associadas a cruzamentos semaforizados selecionados, permitindo que os valores de atraso medidos sejam ligados à estrutura espacial da cidade.

Em torno de cada cruzamento avaliado, são gerados grafos de mobilidade de peões com pontos de teste para simular possíveis movimentações. Estas rotas simuladas permitem relacionar os valores de atraso ao nível da travessia com os indicadores ao nível do cruzamento. Cada cruzamento avaliado é então classificado de acordo com o Nível de Serviço para peões e exportado com métricas de atraso, pontuações de gravidade, informações de alerta e atributos de visualização.

Esta Tese espera obter dois resultados geoespaciais principais. O primeiro é uma camada Geographic JavaScript Object Notation (GeoJSON) ao nível do cruzamento, que distingue entre locais com valores de atraso medidos e cruzamentos semaforizados onde a recolha de dados ainda está incompleta. O segundo é uma camada ao nível da zona, onde as classificações válidas dos cruzamentos são agregadas por área urbana.

Por fim, os resultados mostram que o atraso dos peões pode ser transformado de observações de campo isoladas em informação espacial estruturada para monitorização urbana. Embora a avaliação atual seja limitada pelo pequeno número de valores de atraso medidos manualmente, a abordagem proposta fornece uma estrutura prática e atualizável. À medida que mais medições de campo ou dados baseados em sensores se tornarem disponíveis, a mesma abordagem pode ser expandida para apoiar uma visão mais completa das condições de mobilidade dos peões no Porto e em todo o mundo.

Palavras-chave: mobilidade pedonal • atrasos pedonais • cruzamentos urbanos • Nível de Serviço • dados geoespaciais • OpenStreetMap (OSM) • GeoJSON •

Abstract

Cities increasingly rely on digital platforms to monitor traffic, public transport, air quality, and other urban systems. Yet, one of the most common parts of daily mobility often remains less visible: the pedestrian experience. In particular, waiting time at signalized crossings can interrupt walking routes, reduce the continuity of short trips, and make pedestrian movement less comfortable. Although these delays happen at specific points in the city, they are rarely translated into spatial layers that can be inspected and updated by urban management teams.

This dissertation focuses on signalized urban intersections in Porto, where pedestrian waiting time can directly affect walking conditions. To study this problem, it develops a geospatial workflow that combines spatial data, measured pedestrian delay, route simulation, and delay-based classification. The proposed approach combines open spatial data, field-measured delay values, local pedestrian route simulation, and delay-based Level of Service (LOS) classification. OpenStreetMap (OSM) data is used to extract the road network, identify intersections, detect pedestrian crossings, and distinguish crossings with pedestrian signal control. Field observations are then linked to selected signalized crossings, allowing measured delay values to be connected to the spatial structure of the city.

Around each evaluated intersection, the workflow builds a local pedestrian graph and generates test points to simulate possible pedestrian movements. These simulated routes make it possible to connect crossing-level delay values with intersection-level indicators. Each evaluated intersection is then classified according to pedestrian Level of Service and exported with delay metrics, severity scores, alert information, and visualization attributes.

The proposed workflow produces two main geospatial outputs. The first is an intersection-level Geographic JavaScript Object Notation (GeoJSON) layer, which distinguishes between locations with measured delay values and signalized intersections where data collection is still incomplete. The second is a zone-level layer, where valid intersection classifications are aggregated by urban area.

The results show that pedestrian delay can be transformed from isolated field observations into structured spatial information for urban monitoring. While the current assessment is limited by the small number of manually measured delay values, the proposed approach provides a practical and updatable structure. As more field measurements or sensor-based data become available, the same approach can be expanded to support a more complete view of pedestrian mobility conditions in Porto and all around the world.

Keywords: pedestrian mobility • pedestrian delay • urban intersections • Level of Service (LOS) • geospatial data • OpenStreetMap (OSM) • GeoJSON •

The United Nations Sustainable Development Goals

The United Nations Sustainable Development Goals (SDG) provide a global framework for achieving a better and more sustainable future for all. They include 17 goals that address the world's most pressing challenges, including poverty, inequality, climate change, environmental degradation, peace, and justice.

This dissertation contributes mainly to SDG 11, while also relating to SDG 3.

SDG 3 Ensure healthy lives and promote well-being for all at all ages.

SDG 11 Make cities and human settlements inclusive, safe, resilient and sustainable.

SDG	Target	Contribution	Performance Indicators and Metrics
3	3.6	The work relates indirectly to pedestrian safety and well-being. Although it does not measure accidents or health outcomes, it studies conditions that can influence the walking experience, such as long waiting times at crossings. Understanding these delays can help identify places where pedestrian conditions should be improved to avoid dangerous behaviour like crossing during a red light.	Number of signalized crossings analysed; pedestrian delay values at crossings; number of intersections classified with low pedestrian Level of Service; identification of locations where waiting times may negatively affect pedestrian comfort or crossing behaviour.
11	11.2	This dissertation supports the analysis of pedestrian mobility at urban intersections. By identifying intersections, associating pedestrian crossings, and estimating delay at signalized crossings, the work helps understand where walking conditions may be less comfortable or less efficient. This can support future improvements in pedestrian accessibility and urban planning.	Number of intersections identified; number of pedestrian crossings associated with intersections; number of signalized crossings considered; average pedestrian delay per intersection; maximum pedestrian delay intersection; pedestrian Level of Service classification.
	11.3	The work produces geospatial layers that can be used to help urban analysis and decision-making. Instead of only keeping the results as tables or isolated measurements, the dissertation structures them as spatial information that can be visualized and integrated into a usable tool.	Number of GeoJSON layers generated; number of evaluated intersections with complete attributes; availability of visualization fields such as LOS, severity score and marker colour; number of areas with aggregated pedestrian mobility indicators.

Acknowledgements

Em primeiro lugar, gostaria de agradecer aos meus pais. Obrigado por me terem dado a conhecer a Engenharia Elétrotécnica e por me terem ajudado a descobrir uma área que acabou por marcar profundamente a minha vida académica e pessoal. Mais do que isso, obrigado pelo apoio, paciência e confiança que me demonstraram ao longo destes cinco anos e dos dezoito antes desses.

Gostaria também de agradecer ao resto da minha família, que me acompanhou ao longo desta jornada. Obrigada, Avô Calçada, porque sem as inúmeras boleias acompanhadas de boa companhia não tinha assistido a metade das aulas. Obrigada, Avó Mena e Avó Glória, por estarem sempre mais entusiasmadas que eu para eventos académicos e por me ensinarem tanto sobre tudo.

Aos meus amigos, que estiveram sempre ao meu lado a sofrer as mesmas dores, obrigado. Vou sempre recordar com carinho todas as memórias que fizemos, desde as noites a fazer diretas nos anfiteatros a tentar ensinar uns aos outros matéria de vários meses, até todos aos FEUPcaffes e tardes em que devíamos estar a estudar, mas acabávamos por aproveitar o convívio na AE. A vossa presença, o vosso incentivo e a vossa boa disposição tornaram os momentos mais difíceis mais fáceis e os bons momentos mais memoráveis. Um especial obrigada à Matilde e ao Jota por serem os meus verdadeiros colegas de curso. Obrigada às Ranços por serem as mais divertidas. E claro, gostaria também de deixar um agradecimento muito especial ao Vasco, o meu maior apoio. Sem ti, tudo isto teria sido muito menos divertido.

Expresso também a minha gratidão ao meu orientador, o Professor Daniel G. Costa, por ter aceite orientar esta dissertação e por me ter dado todo o apoio de que precisei ao longo do processo. A sua orientação, disponibilidade e feedback foram essenciais para o desenvolvimento deste trabalho.

Por fim, gostaria de agradecer a toda a equipa do Porto Digital por me ter acolhido e acompanhado ao longo deste processo. Estou especialmente agradecida à equipa de Dados e IA, ao Henish Balu e ao Miguel Almeida pela orientação e apoio. Gostaria de deixar uma menção especial de agradecimento ao Miguel Almeida, cujo empenho no sucesso deste projeto e as suas constantes contribuições fizeram uma diferença significativa no resultado desta dissertação.

A todos os que participaram nesta aventura, obrigada.

Maria Calçada

Declaration of honour

I declare that this dissertation is my own work and has not previously been Submitted or assessed at this or any other institution. References to other authors (statements, ideas, thoughts), as well as the use of published works, strictly comply with the rules of attribution and are duly identified in the text and in the bibliographic references, in accordance with the applicable referencing standards. I am aware that the practice of plagiarism or self-plagiarism constitutes an academic offence, as established in the U.Porto Code of Ethics.

I also declare that I have used Artificial Intelligence tools to complete part(s) of this dissertation, and that this use and its results are duly noted and have been carefully checked for errors, inaccuracies, unfounded attributions or other forms of bias in content and arguments, as well as determining authorship.

Maria Pinto Calçada

Maria C.

"There is no logic that can be superimposed on the city; people make it, and it is to them, not buildings, that we must fit our plans."

Jane Jacobs

Contents

List of Figures

List of Tables

List of Acronyms

AI Artificial Intelligence
APD Associação Porto Digital
BLE Bluetooth Low Energy
CSV Comma-Separated Values
EPSG European Petroleum Survey Group
GDPR General Data Protection Regulation
GeoJSON Geographic JavaScript Object Notation
IoT Internet of Things
ITS Intelligent Transportation Systems
LOS Level of Service
OSM OpenStreetMap
OSMnx OpenStreetMap Network Analysis Tool
SDG Sustainable Development Goals
WGS84 World Geodetic System 1984

Chapter 1

Introduction

Walking is one of the most basic forms of urban mobility, but it is also one of the easiest to overlook in data-driven urban management. In many cities, road traffic, public transport, parking, air quality and other urban systems are already well represented through digital layers and monitoring tools [1], [2], [3], [4]. Pedestrian mobility, however, is often harder to describe in the same way, especially when the focus is not only on where people can walk, but also on how comfortable and continuous that movement is [5], [6], [7].

This dissertation focuses on pedestrian conditions at urban intersections, with particular attention to signalized pedestrian crossings. These locations are important because they interrupt walking routes and introduce waiting time into pedestrian movement [8], [9], [10], [11]. A pedestrian may have a short route in terms of distance, but that route can still feel inconvenient if it includes several crossings with long red phases or poorly coordinated pedestrian signals. For this reason, pedestrian delay is used in this work as a measurable indicator of how signalized crossings affect local walking conditions.

A simple example of this issue is shown in Figure 1.1. A walking trip is often represented as movement from one intersection to another, but the actual experience also depends on what happens at each crossing. A pedestrian may lose time waiting at one signalized crossing, then again at the next one. In a short walking trip of five to ten minutes, these repeated delays can become a relevant part of the total travel time. This motivates the need to treat intersections not only as geometric points in the street network, but also as locations where pedestrian movement can be interrupted.

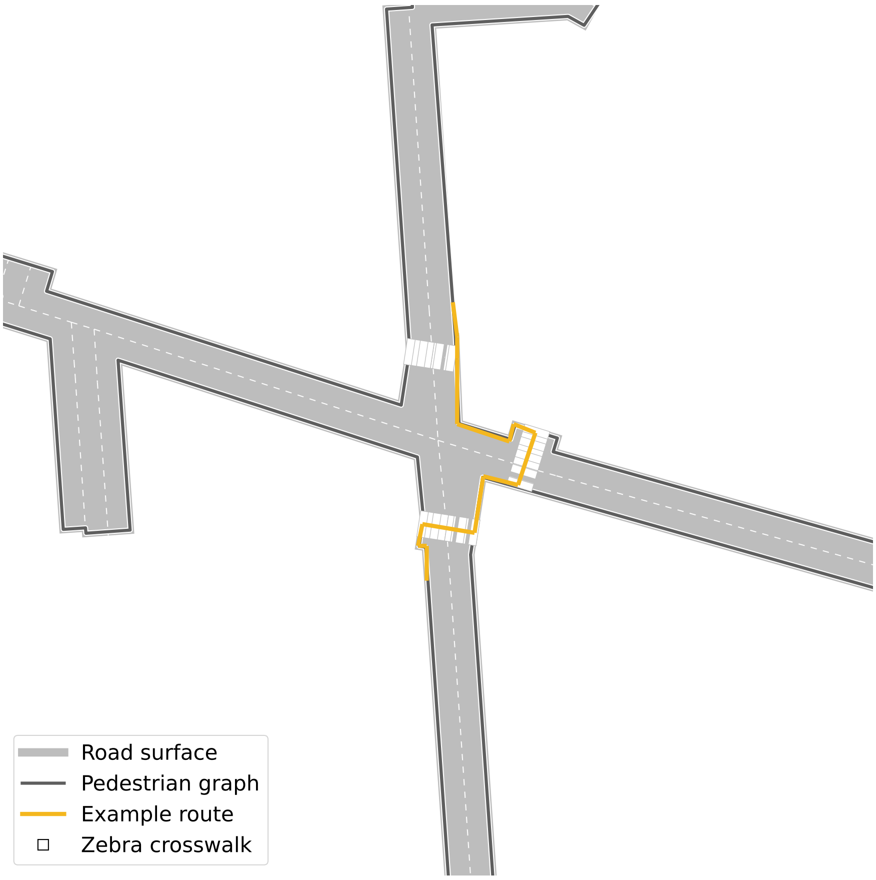

Figure 1.1: Illustrative example of a local pedestrian route around a signalized intersection. The route shows how crossing delays can become part of the total walking time, motivating the evaluation of pedestrian conditions at intersections. Source: Author

As a central case study, this work is developed in the context of the municipality of Porto. Porto is a compact and dense city, where short-distance movement, access to public transport and interaction between pedestrians and road traffic are part of daily urban mobility. In this setting, intersections and crossings play an important role in the walking experience. Understanding where pedestrian delay occurs and how it can be represented spatially can support a clearer view of pedestrian mobility conditions across the city.

At the same time, Porto has been developing digital urban management initiatives, including the use of data platforms to organize, visualize and monitor city information [12], [13]. In this dissertation, the term urban monitoring platform refers to a digital environment used by city teams to visualize, inspect and follow urban information in a spatial and operational way. It is understood as a practical interface built on top of urban data sources, where geospatial layers and indicators can be displayed to support interpretation and decision-making. This creates a relevant framework for the present work, since the goal is not only to calculate pedestrian delay indicators, but also to structure them as spatial layers that can be inspected, updated and reused in this type of environment.

This type of structure may also become relevant in the context of future large-scale urban events. Portugal is one of the host countries of the 2030 FIFA World Cup, and Porto is expected to be one of the Portuguese host cities. Events of this scale can change urban mobility dynamics, especially around public transport access, pedestrian flows and movement between busy areas. Although this dissertation does not apply the proposed workflow directly to that event, it creates a spatial structure that could support future studies and planning tasks related to pedestrian mobility during periods of higher urban demand.

The final result of the dissertation is therefore intended to be useful as both an analytical output and a possible urban data layer with real-world applicability.

1.1 Motivation and Context

Pedestrian mobility is essential for the functioning of cities. Even when trips are made mainly by car, bus, metro, or another transport mode, they usually include at least a short walking segment. Walking connects people to public transport stops, local services, schools, workplaces and public spaces. Because of this, the quality of pedestrian movement affects not only those who walk as their main mode of transport, but also the wider mobility system [14], [15], [16].

However, pedestrian mobility is not only about the existence of sidewalks or crossings. A route can be physically possible and still be uncomfortable, slow, or interrupted. This is especially visible at urban intersections, where pedestrians often need to wait, change direction, cross in more than one stage, or interact with vehicle traffic. In dense urban areas, these interruptions can occur repeatedly within short distances.

Signalized pedestrian crossings are a clear example of this problem. Traffic signals are necessary to manage conflicts between pedestrians and vehicles, but they can also create delay. Long waiting times can reduce the continuity of walking routes and make pedestrian movement less attractive. In some cases, excessive waiting may also encourage unsafe behaviour, such as crossing before the pedestrian signal turns green.

Despite this, pedestrian delay is not always represented in urban monitoring tools. Cities may have spatial information about streets, crossings, traffic signals, or public transport, but this information is often not organized as a direct assessment of pedestrian conditions [1], [2], [3]. Field measurements can provide useful information about waiting times [8], [9], [10], but they are usually limited to a small number of locations and need to be connected to a spatial structure before they can support broader analysis.

This dissertation is motivated by that gap, exploring how open geospatial data, field-measured delay values and route simulation can be combined to assess pedestrian conditions at signalized intersections. The aim is to move from scattered spatial elements and manual observations to structured outputs that can be visualized at intersection level and aggregated by urban zone.

The motivation is also practical. By exporting the results as Geographic JavaScript Object Notation (GeoJSON) layers, the assessment can be represented in the urban monitoring platform used by the city of Porto [12], [17]. This makes the results easier to inspect and creates a structure that can be expanded later as more delay measurements or sensor-based data become available.

1.2 Problem Statement

Although pedestrian mobility is an important part of urban mobility, it is not always easy to evaluate it in a way that can be used by urban management platforms. Pedestrian conditions depend on several local factors, including crossings, signal timings, waiting times and the way routes pass through intersections. These elements are spatial, but they are often stored in different formats or are not directly available as ready-to-use assessment layers.

Open geospatial data, such as OpenStreetMap (OSM), provides useful information about road networks, crossings and traffic signals [18], [19], [20], [21]. However, this data needs to be processed before it can support a pedestrian delay assessment. Intersections need to be identified, crossings need to be extracted and associated with them and signalized pedestrian crossings need to be distinguished from other crossings. Without this processing, the available spatial data remains too fragmented for the type of evaluation developed in this work.

Another difficulty is the limited availability of pedestrian delay data. Waiting times at crossings can be measured in the field, but this process is manual and difficult to scale to a large number of locations. As a result, measured delay values are often available only for a small subset of intersections. This creates the need for a structure that can use the available measurements while also keeping the remaining locations prepared for future data collection.

The problem addressed in this dissertation is therefore the lack of a structured data processing workflow that connects open spatial data, manually measured pedestrian delay, route simulation and platform-ready geospatial outputs. This work therefore aims to show how these elements can be combined to assess signalized intersections and produce spatial layers that are understandable, updatable and useful for urban monitoring.

1.3 Research Question and Objectives

Based on the problem described above, this dissertation is guided by the following research question:

> How can geospatial data and pedestrian delay be combined to assess pedestrian conditions at signalized intersections and represent the results in an urban monitoring platform?

To answer this question, the work defines the following objectives:

- Extract and prepare the road network, intersections, pedestrian crossings and traffic signal information from open geospatial data.
- Identify the intersections where pedestrian delay can be evaluated, focusing on locations associated with signalized pedestrian crossings.
- Integrate manually measured delay values collected at selected pedestrian crossings.

- Simulate local pedestrian routes around intersections in order to connect crossing-level delay values with intersection-level movement.
- Calculate delay-based indicators and classify intersections according to pedestrian Level of Service (LOS).
- Export the results as geospatial layers that can be visualized at intersection level and aggregated by urban zone.
- Prepare the final outputs in a format suitable for use in the urban monitoring platform used by the city of Porto.

1.4 Contributions

The main contribution of this dissertation is the development of a geospatial workflow for assessing pedestrian conditions at signalized urban intersections. The workflow combines open spatial data, manually measured delay values, route simulation and delay-based classification to produce results that can be visualized and reused in an urban monitoring context.

More specifically, the work contributes with:

- A reproducible process for extracting and preparing road network data, intersections, pedestrian crossings and traffic signal information from OSM.
- A method for associating pedestrian crossings with nearby intersections and identifying the subset of intersections where signalized pedestrian delay can be evaluated.
- The integration of manually measured pedestrian delay values into a spatial workflow, allowing field observations to be linked to specific crossings and intersections.
- A local pedestrian route simulation approach that uses generated test points and a pedestrian graph to connect crossing-level delays with possible movements around each intersection.
- A delay-based LOS classification for evaluated intersections, including severity scores, alert indicators and visualization attributes.
- The production of final GeoJSON layers at intersection level and zone level, prepared for visualization in the urban monitoring platform used by the city of Porto.

These contributions are intended to support both the current assessment and future extensions of the work. Although the number of measured delay values is limited, the workflow defines a structure that can be updated as more field measurements or sensor-based data become available.

1.5 Dissertation Structure

This dissertation is organized into six chapters.

Chapter 1 introduces the motivation, problem statement, research question, objectives and main contributions of the work.

Chapter 2 presents the background and state of the art. It discusses pedestrian mobility, walkability, pedestrian delay, LOS, geospatial data, graph-based models, urban data platforms and possible future sensing technologies.

Chapter 3 describes the methodology developed in this dissertation. It explains the study area, data sources, intersection and crossing extraction, measured delay integration, route simulation, LOS classification, GeoJSON export and area-level aggregation.

Chapter 4 presents the technical implementation of the methodology. It describes the development environment, main tools, processing stages, parameters, algorithms, intermediate files and final outputs generated by the workflow.

Chapter 5 presents and discusses the results. It focuses on the final point-based and area-level outputs as visualized in the urban monitoring platform, explains their attributes and interpretation and discusses how the work connects with smart city systems and future sensor-based data sources.

Chapter 6 concludes the dissertation. It summarizes the work developed, highlights the main outcomes, discusses the main limitations and future work and closes with final remarks on the contribution of the dissertation.

Chapter 2

Background and State of the Art

This chapter introduces the main concepts, methods and technologies that support the work developed in this dissertation. It provides the context needed to understand why pedestrian delay can be studied through geospatial data, network analysis and spatial tools.

2.1 Pedestrian Mobility and Walkability

Pedestrian mobility is one of the basic components of urban mobility. Even when a trip is mainly made by car, bus, metro, or another transport mode, it usually begins or ends with walking. For this reason, the quality of pedestrian movement affects short walking trips, access to public transport, local services, schools, workplaces and public spaces.

Walking is one of the most accessible and low-impact modes of urban mobility. It does not require fuel, produces no direct emissions during use and can support healthier daily routines. These qualities are also reflected in proximity-based planning concepts such as the 15-minute city, which proposes that essential daily services should be accessible within a short walk or bicycle trip from home, reducing the need for long and frequent motorized trips [22]. Recent empirical work on Barcelona also suggests that proximity to everyday destinations can reduce car use mainly by influencing modal choice, rather than by reducing the driving time of those who continue to use the car [16].

Walking is therefore associated with several public benefits, which are summarized in Table 2.1. These benefits help explain why pedestrian mobility should be considered in urban planning and why cities need infrastructure that makes walking safe, continuous, comfortable and practical.

Table 2.1: Main benefits associated with walking.

Benefit	**Description**	**References**
Public health	Walking supports daily physical activity, helps reduce sedentary behaviour and can contribute to mental well-being through contact with public space and everyday urban life.	[14], [15]
Environmental sustainability	Walking is a low-impact transport mode. It produces no direct emissions during use and can replace some short motorized trips when destinations are accessible on foot.	[6], [22], [23]
Reduction of car traffic	When daily services and activities can be reached on foot, some trips that would otherwise be made by car can be avoided. This can reduce car use and help relieve pressure on urban road networks.	[16], [22], [24]

Still, walking does not automatically become attractive just because it is possible. A city can have sidewalks and crossings and still offer a poor walking experience if pedestrian routes are interrupted, unsafe, uncomfortable, exposed to weather conditions, or difficult to follow. This is where the concept of walkability becomes relevant.

The concept of walkability is commonly used to describe how suitable an urban environment is for walking. It usually includes aspects such as safety, comfort, accessibility, connectivity, land-use diversity, protection from weather conditions and the continuity of pedestrian routes. This protection is relevant because exposure to rain, excessive sun, or lack of shade can affect how comfortable and attractive walking is in daily routines. Dovey and Pafka describe walkability through the relationship between density, mix and access, showing that walking conditions depend on more than the physical presence of pedestrian infrastructure [5]. Other work on walkability assessment also shows that the same urban environment can be experienced differently by different pedestrian groups [6], [7].

Recent work with case studies in Bologna and Porto also shows that perceived walkability depends on several local urban characteristics, including accessibility, comfort and the way people experience the surrounding environment [25]. This is relevant for this dissertation because Porto is not treated only as a geometric network. It is an urban setting where walking conditions depend on the way streets, crossings and signalized locations are experienced in practice.

These studies are useful because they show that pedestrian mobility is affected by many urban factors. They also make clear that walkability can become difficult to measure in a direct and operational way. If every aspect of the walking experience is included at the same time, the assessment may become too broad for a practical urban management tool. This is one of the reasons why several data-driven approaches use more specific indicators, such as delay, connectivity, accessibility, or route continuity.

Data-driven approaches have become more common in pedestrian mobility studies. They make it possible to move from general descriptions of the urban environment to indicators that can be measured and compared. Costa et al. highlight the role of geospatial data in supporting smart and sustainable urban analysis, especially when different types of urban information need to be combined in a spatial framework [1]. Guo et al. also show how pedestrian-oriented indicators can be used to diagnose the quality of streets and support more detailed assessments of walking conditions [26]. Other recent work also explores the relationship between urban form and pedestrian activity through data-driven models, reinforcing the importance of measurable indicators in walkability analysis [27].

Urban intersections are especially relevant in this context. Even when the pedestrian network between intersections is adequate, crossings can introduce delays and interruptions that affect the quality of a walking route. This is common in dense urban areas, where pedestrians may need to cross several streets within a short distance. Repeated waiting times can affect the continuity of the route and may make walking feel less convenient.

Signalized crossings are a clear example of this issue. Traffic lights are needed to manage conflicts between pedestrians and vehicles, but their timing can strongly influence pedestrian conditions. Long red phases, short green phases, or poorly coordinated pedestrian signals may increase waiting times and reduce the continuity of walking routes. Previous studies point out that waiting time and delay are relevant parts of the pedestrian experience, especially at intersections where movement is controlled by traffic signals [9], [10], [26].

In Porto, these issues are particularly relevant. The city has a concentrated and irregular urban fabric, with narrow streets, frequent intersections, changes in elevation and many areas where pedestrians, vehicles and public transport share limited space. In this context, the quality of pedestrian mobility depends on how intersections and crossings work in practice.

For this dissertation, walkability is not treated as a full urban quality index. The work focuses on a narrower part of the pedestrian experience: intersections, crossings and delay. This makes the problem more measurable and connects directly to the proposed methodology, which extracts intersections and crossings from geospatial data, integrates observed delay values, simulates local pedestrian routes and classifies the evaluated intersections according to pedestrian LOS.

2.2 Urban Intersections and Pedestrian Delay

Urban intersections are points where different movements meet. For pedestrians, they are often the places where a walking route is interrupted. A person may need to wait for a green signal, check for vehicles, change direction, or cross more than one road segment before continuing the route. Because of this, pedestrian mobility cannot be understood only through sidewalks or street connectivity. Intersections also need to be considered.

At signalized intersections, traffic signals regulate how pedestrians and vehicles share the same space. This regulation is needed for safety, especially in busy areas, but it also introduces waiting time into pedestrian routes. When pedestrian green phases are short, red phases are long, or crossings are divided into several stages, the time spent waiting can become a relevant part of the total walking time.

Figure 2.1 illustrates the type of urban situation addressed in this dissertation. At signalized crossings, pedestrians may need to wait before continuing their route and this waiting time becomes part of the walking experience. The example also shows how different users share the same crossing environment, reinforcing the relevance of studying delay from a pedestrian perspective.

Figure 2.1: Field observation of pedestrians waiting at a signalized crossing in Porto during the delay measurement campaign. Source: Author

Pedestrian delay is usually understood as the additional time pedestrians spend waiting at a crossing compared with a situation where they could cross without stopping. This delay is commonly used to assess the performance of signalized crossings from the pedestrian point of view, because it is simple to measure and directly connected to the walking experience [8], [10], [11]. In this dissertation, delay is used as the main variable for evaluating how signalized crossings affect local pedestrian routes.

The importance of pedestrian delay is also related to behaviour. Long waiting times can make walking less pleasant, especially when they occur several times along the same route. Previous studies suggest that waiting time and crossing conditions can influence pedestrian behaviour at intersections, including crossing decisions and interaction with vehicles [9], [28]. This connects delay with travel time, comfort, perceived safety and compliance.

The effect of delay can be stronger in dense urban environments. In areas with many intersections close to each other, a pedestrian may experience several interruptions within a short walking distance. Even if each delay is moderate, the combined effect can reduce the continuity of the route. This helps explain why intersections are treated as a central element in this dissertation.

A limitation in many pedestrian delay studies is that the analysis is often focused on individual intersections or isolated crossings. This type of analysis is useful for detailed observation, but it does not always show how different intersections compare across a wider area. It also does not automatically produce spatial outputs that can be used by urban management platforms.

This dissertation addresses that limitation by analysing pedestrian delay at the level of crossings and intersections, while keeping the results in a geospatial structure. Field observations are used to introduce measured delay values for selected signalized crossings. These values are then associated with local pedestrian routes and with the corresponding intersections. This allows the methodology to compare intersections through delay-based indicators, instead of describing them only through their location or geometry.

2.3 Pedestrian Level of Service

Pedestrian LOS is commonly used to describe pedestrian conditions in a way that is easier to read than raw numerical values. Instead of presenting only waiting times or delay values, LOS converts performance into classes. This helps compare different locations and makes the results easier to communicate.

In transport studies, LOS is often used to evaluate how well a transport facility performs from the user's point of view. For pedestrians, this evaluation may include available space, comfort, safety, crossing conditions and delay. In the case of signalized crossings, pedestrian delay is one of the most direct indicators, because it measures the time pedestrians spend waiting before they can continue their route [8], [10].

Delay-based LOS is useful because it is simple to interpret. A low delay usually means that the pedestrian can cross with little interruption. A high delay indicates that the crossing adds a stronger time penalty to the route. This does not describe the full walking experience, but it captures an important part of it in places where traffic signals control pedestrian movement.

Several studies use pedestrian delay as an input to evaluate conditions at signalized intersections. Guo et al. discuss delay as part of the diagnosis of pedestrian street quality, while Papandreou et al. show how signal timing can affect pedestrian LOS at signalized intersections [26], [29]. Other studies also propose LOS methods for pedestrian crossings and signalized intersections, considering aspects such as crossing conditions, user perception and pedestrian characteristics [11], [30], [31].

There are also more detailed LOS approaches that include several aspects of the crossing experience. These may consider comfort, traffic conditions, perceived safety, or pedestrian characteristics [32]. Such approaches can provide a richer description, but they also require more data and may be harder to apply consistently across many intersections.

For this reason, LOS needs to be interpreted with care. Pedestrian experience is affected by more than delay. A crossing with a short waiting time may still feel uncomfortable if the road is wide, traffic is fast, visibility is poor, or the waiting environment feels unsafe. Recent work on pedestrian waiting experiences also shows that the perception of waiting can depend on the surrounding context and on the way the crossing is designed [9]. LOS is therefore useful as an evaluation tool, but it should not be read as a complete measure of walkability.

Waiting time can also influence crossing behaviour. When delays are perceived as too long, pedestrians may be more likely to cross during the red phase, which can increase conflicts with vehicles [28]. This reinforces the relevance of delay as an indicator for the type of assessment developed in this dissertation.

In this work, LOS is used in a focused and transparent way. The aim is to classify intersections according to the delay introduced by signalized crossings. This choice makes the evaluation easier to reproduce and easier to connect with the geospatial workflow. Each evaluated intersection can then be represented with a LOS class, a severity score and an associated visualization colour.

This classification is useful for integration into an urban management platform. A map with delay values alone may be harder to interpret when many intersections are shown at the same time. By assigning each evaluated intersection to a LOS class, the output becomes easier to compare and can help identify locations where pedestrian delay deserves closer attention.

2.4 Geospatial Data for Pedestrian Mobility Analysis

Geospatial data is useful in pedestrian mobility studies because walking is strongly tied to place. Streets, crossings, intersections, traffic signals, sidewalks and barriers all have a location and their position affects how people move through the city. Analysing pedestrian mobility only through general indicators can therefore miss important spatial details. A crossing that creates delay affects a specific point of the pedestrian network and can influence the routes that pass through it.

OSM has become a common source for this type of analysis because it provides open information about urban networks. It can include roads, pedestrian paths, crossings, traffic signals and other elements that are relevant for mobility studies. Its main advantage is that it allows researchers to build spatial datasets without starting from zero. Bartzokas-Tsiompras, for example, shows how OSM data can be used to measure pedestrian street infrastructure at a large scale [20].

Still, OSM data needs to be handled carefully. The same type of pedestrian element may be mapped in different ways depending on the contributor, the city, or the level of detail available. A crossing may appear as a point, as a line, or may be missing. Traffic signals can also be represented with different tags or levels of precision. Recent work on signalized intersections using OSM also shows that intersection information often needs to be inferred, completed or structured before it can be used in transport analysis [21]. Using OSM for pedestrian analysis therefore requires filtering, cleaning and decisions about how each element should be interpreted.

This limitation is directly related to the work developed in this dissertation. The methodology depends on identifying intersections, extracting pedestrian crossings, distinguishing crossings with and without pedestrian signals and associating each crossing with a nearby intersection. These steps are spatial by nature. They require distances, buffers, projections and geometric operations, especially because visualization coordinates and metric calculations do not use the same coordinate system.

Geospatial data also makes it possible to connect field measurements with a wider spatial structure. Field observations are important, but they usually cover a limited number of locations. When these observations are linked to spatial layers, they can be placed in context and compared with other intersections. This helps understand where delay was measured and how each measured crossing relates to the surrounding pedestrian network.

Costa et al. discuss the role of geospatial data-driven approaches in smart city analysis, especially when different urban data sources need to be combined to support decision-making [1]. In this dissertation, this idea is applied to pedestrian mobility. Open geospatial data provides the base structure, while measured delay values add information about how selected crossings work in practice.

The final output of this process is exported as GeoJSON. This format stores geometry and attributes in the same file, which allows each evaluated intersection to keep its location together with values such as average delay, maximum delay, LOS, severity score and visualization colour [17]. This structure makes the results easier to inspect on a map and prepares the output for future integration into the urban monitoring platform used by the city of Porto.

Existing geospatial approaches provide a useful base for pedestrian analysis, but the data is rarely ready to be used directly for intersection-level delay assessment. This dissertation addresses that gap by turning OSM-based information into a structured workflow for extracting intersections and crossings, associating measured delays and preparing a final layer for visualization and decision support. The next section continues from this point by looking at network-based and graph models, including tools such as OpenStreetMap Network Analysis Tool (OSMnx) [18].

2.5 Network-Based and Graph Models

Network-based models are often used to represent movement in urban areas. In these models, the city is represented as a graph, where nodes can correspond to intersections, crossings, or decision points and edges correspond to the connections between them. This structure makes it possible to compute routes, analyse connectivity and estimate travel times across the network.

For pedestrian mobility, this representation is useful, but it needs to be used with care. Pedestrian movement does not always follow the same logic as vehicle movement. Pedestrians may use crossings, sidewalks, stairs, plazas, informal paths and other spaces that are not always represented in road network models. Jabbari et al. discuss this distinction and show that pedestrian networks have their own structure, which should be considered when modelling pedestrian movement [33].

OSM has made this type of analysis easier because it provides spatial information that can be converted into network data. OSMnx is one of the main tools used for this purpose. It allows users to download, study, build and visualize street networks from OSM in a reproducible way [18]. More recent work also shows how OSMnx can be used to analyse urban networks and related amenities, making it relevant for different types of urban studies [19].

When studying pedestrian routes, graph models are useful because they allow us to understand how different parts of the urban network are connected. They can support the calculation of shortest paths, route distances, accessibility measures and network-based indicators. Dhanani et al. show how pedestrian demand and network structure can be considered in active transport evaluation and planning [34]. This type of work is useful when the goal is to move from isolated observations to a more structured view of pedestrian movement.

However, graph-based models can also simplify the pedestrian experience too much. If a model only considers distance or connectivity, it may treat movement through the network as continuous. In real walking routes, this is often not the case. Many elements such as crossings, traffic signals, waiting times, stairs, barriers or local street conditions can change the performance of a route. This is especially relevant at intersections, where pedestrians may need to wait before continuing or cross in more than one stage.

This limitation is relevant for the work developed in this dissertation. A route with a short distance may still be affected by long waiting times and two intersections with similar geometry may perform differently because of their signal timings. For this reason, network-based modelling needs to be combined with crossing-level delay information when the objective is to evaluate pedestrian conditions at signalized intersections.

Simulation can also support this type of analysis when direct observations are limited. Recent work on pedestrian road-crossing simulation, for example, shows how simulated crossing behaviour can be connected with dynamic traffic light scheduling systems [35].

In this dissertation, graph-based modelling is used to simulate local pedestrian routes around intersections. The network provides the spatial structure for movement and the measured delay values represent the time penalty associated with signalized crossings. This makes it possible to classify evaluated intersections using spatial information and delay-based indicators.

2.6 Urban Data Platforms and Decision Support

Urban mobility analysis is increasingly connected to digital platforms that collect, organize and visualize different types of city data. These platforms can support decision-making by bringing together information from several sources, such as traffic sensors, public transport systems, environmental monitoring and geospatial datasets. In this context, data only becomes useful when it can be represented clearly and used by the people who need it.

Smart city research often highlights the role of data-driven methods in urban management. Costa et al. discuss how geospatial data can support smart city applications by helping cities understand urban systems through spatial information [1]. Goumiri et al. also describe smart mobility as an area where data, digital tools and transport systems are increasingly connected [2]. Urban data platforms are also discussed as interfaces for smart governance and as tools that raise practical questions about data organization, reuse and decision-making [3], [4]. These approaches are relevant for this dissertation because the final result is intended to be more than a static analysis. The final result is structured as a spatial output that can be visualized and reused.

For pedestrian mobility, this type of integration is still less common than for vehicular traffic. Many cities already monitor car traffic through sensors, traffic counts, travel time data, or congestion indicators. Pedestrian mobility is harder to capture in the same way, especially when the objective is to understand conditions at crossings and intersections. As a result, pedestrian indicators are often less present in operational platforms, even when walking is an important part of daily mobility in a city.

This creates a gap between pedestrian mobility research and practical urban management. Studies may calculate indicators such as walkability, delay, or accessibility, but these results are not always transformed into formats that can be used by city teams. A table of delay values may be useful for analysis, but it is harder to interpret operationally without a spatial representation. For this reason, the way results are exported and visualized becomes part of the usefulness of the methodology.

GeoJSON is useful in this context because it is a widely used format for representing geospatial data and allows geographic features and their attributes to be stored together [17]. In this dissertation, each evaluated intersection can be represented as a point with information such as average delay, maximum delay, LOS, severity score and visualization colour. This structure helps address a gap in operational urban tools, where pedestrian-specific indicators are often less present than vehicular traffic indicators. It also makes the output suitable for integration into urban data platforms and for visualization in urban monitoring environments used by city teams.

2.7 Sensing Technologies and Future Data Sources

Pedestrian mobility can be studied with different types of data. Field observations are useful because they provide direct information about what happens at a specific crossing, such as pedestrian green time, red time, or observed waiting time. However, collecting this type of data manually is slow and difficult to repeat across a large number of intersections. For this reason, sensing technologies are often discussed as a possible way to collect mobility data more continuously.

Table 2.2: Sensing technologies and possible future data sources for pedestrian mobility assessment.

Technology / data source	**Potential contribution**	**Main limitations**	**References**
Field observations	Direct measurement of pedestrian signal timings, crossing conditions and waiting times at specific locations.	Slow to collect manually and difficult to scale to a large number of intersections.	[10], [11]
Bluetooth monitoring	Estimation of travel times and movement patterns through the detection of devices moving through the network.	May not represent all pedestrians, depends on device detection rates and raises privacy and data quality concerns.	[36], [37]
Wi-Fi and Bluetooth Low Energy (BLE) sensing	Possible estimation of pedestrian presence, movement, dwell time, densities, waiting times and flows using wireless signals emitted by devices.	Coverage and representativeness depend on device behaviour, sensor placement, detection rates and privacy safeguards.	[36], [37]
Mobile network data from telecommunication operators	Estimation of aggregated presence and movement patterns at municipal scale. This can support a broader view of entries, exits, movement patterns and densely occupied areas across the city.	The spatial resolution is usually too coarse for intersection-level or crossing-level analysis. It may not distinguish pedestrians from other mobile users and depends on operator coverage, data processing methods, privacy safeguards and commercial agreements.	[38]
Camera-based sensing	Collection of pedestrian counts, trajectories, crossing behaviour and pedestrian–vehicle interaction data through video analysis at intersections.	Requires image processing, careful sensor placement and strong privacy protection. Performance may depend on lighting, occlusion, weather conditions and camera angle.	[28], [39], [40]
Adaptive signal control	Use of real-time data to adjust signal timings according to observed traffic conditions. In a pedestrian context, this could support more responsive crossing times.	Requires reliable real-time inputs and integration with signal control systems. Existing approaches often focus mainly on vehicles.	[29], [35], [41]

Traffic management systems already use sensor data to support signal operation. Physical detectors, such as inductive loops, can be used at traffic signals to detect vehicle presence and queues, supporting actuated signal control [42]. Although these technologies are mainly designed for vehicular traffic, they show how sensing infrastructure can be connected to signal timing and urban traffic management.

As shown in Table 2.2, pedestrian-oriented sensing can be supported by several types of data sources. Wireless approaches can help estimate movement patterns, presence and dwell time, while camera-based systems can provide more detailed information about trajectories and crossing behaviour [28], [36], [37], [39]. However, these methods still depend on sensor placement, data quality and privacy safeguards, which makes their use more complex than manual field observations.

Adaptive traffic signal control is also relevant for future pedestrian mobility applications. Real-time data can be used to adjust signal timings according to observed traffic conditions. Previous work shows that dynamic signal timing can improve pedestrian LOS at signalized intersections [29]. More recent approaches also consider pedestrian queues in traffic signal control, which is relevant when pedestrian demand should influence signal operation more directly [41]. Studies on alternative signalized intersection designs further show that pedestrian safety and efficiency depend on how signal phases and crossing movements are organized [43].

Because these technologies are not necessarily available across all intersections, this dissertation does not depend on continuous sensor data. The proposed workflow is built using open geospatial data and field-measured delay values, making it easier to apply in a context where detailed pedestrian sensing data is not yet available. At the same time, the output structure leaves room for future data sources. If sensor-based pedestrian counts, real-time waiting times, or signal timing data become available, they could be added to the same geospatial framework.

Future sensor-based extensions would also need to consider privacy and data protection. Pedestrian mobility data can reveal movement patterns, routines or the presence of people in specific places. This is especially relevant for wireless sensing, camera-based systems and smartphone or wearable data. In the European context, these systems would need to respect General Data Protection Regulation (GDPR) principles, including data minimization, purpose limitation and appropriate safeguards for personal data [44]. For this reason, approaches such as aggregation, anonymization, local processing and limited data retention would be important in any future implementation.

The current assessment is therefore static and reproducible, while future sensing technologies could make it more dynamic. In that scenario, the same type of spatial layer could be updated with more frequent measurements and used to support pedestrian-oriented decisions in an urban management platform.

2.8 Identified Research Gaps

The previous sections show that pedestrian mobility has been studied from several points of view, including walkability, pedestrian delay, LOS, geospatial analysis, graph-based modelling, urban data platforms and sensing technologies. These topics provide a useful base for this dissertation, but they also reveal some limitations that motivate the work developed here.

The first limitation is related to the broad nature of walkability. Walkability studies often include several dimensions, such as safety, comfort, accessibility, land-use diversity, street design and pedestrian perception. This gives a rich view of pedestrian mobility, but it can be difficult to translate directly into an operational layer for city management. For this reason, this dissertation focuses on a more specific part of the pedestrian experience: delay at crossings and intersections.

A second limitation concerns the scale at which pedestrian delay is often studied. Many studies analyse delay at individual crossings or selected intersections. This type of analysis is useful for understanding local conditions, but it does not always produce a spatial structure that allows several intersections to be compared across a wider area. In the context of this dissertation, this gap is addressed by associating measured delay values with geospatial intersection data.

A third limitation is related to the use of LOS. LOS is useful because it converts numerical values into classes that are easier to interpret. However, more detailed LOS models may require data that is not always available, such as pedestrian volumes, traffic volumes, comfort perception, or detailed geometric conditions [30], [31], [32]. This dissertation therefore adopts a delay-based LOS classification. The choice is simpler, but it is transparent and compatible with the available data.

Open geospatial data also presents some challenges. OSM provides a strong base for urban mobility analysis, but pedestrian elements are not always mapped in a consistent way. Crossings may appear as points or lines and traffic signals may have different tags or levels of detail. This creates the need for a processing workflow that can clean, interpret and connect these elements before they are used for analysis [20], [21].

Graph-based models introduce another limitation. They are useful for representing connectivity and computing routes, but they often treat movement through the network as continuous. For pedestrians, this can be misleading at signalized intersections, where waiting time may represent a relevant part of the route. This dissertation addresses this by combining local route simulation with delay values associated with signalized crossings.

Urban data platforms create a final motivation for the work. Many platforms already include indicators related to vehicular traffic, environmental conditions, or public transport. Pedestrian indicators are less common, especially at the level of crossings and intersections. There is therefore a need to structure pedestrian mobility results in formats that can be visualized and reused by urban management tools [3], [4].

Based on these gaps, this dissertation develops a geospatial workflow for pedestrian intersection assessment. The workflow extracts intersections and crossings from open spatial data, integrates field-measured delay values, simulates local pedestrian routes, assigns delay-based LOS classes and exports the results as GeoJSON layers. The goal is to create an output that can support both analysis and future integration into urban management platforms.

2.9 Summary

This chapter reviewed the main concepts and previous work that support the dissertation. It started by discussing pedestrian mobility and walkability, showing that walking conditions depend on several factors, including accessibility, comfort, connectivity and the continuity of pedestrian routes. Since walkability is a broad concept, the chapter then focused on a more specific part of the pedestrian experience: the role of intersections and delays at signalized crossings.

Pedestrian delay and LOS were presented as useful indicators for evaluating crossing conditions. The literature shows that waiting time can affect travel time, comfort and pedestrian behaviour. At the same time, LOS was discussed as a simplified way of translating delay values into classes that are easier to compare and visualize.

The chapter also reviewed the use of geospatial data, OSM and graph-based models in pedestrian mobility analysis. These tools make it possible to represent intersections, crossings and routes in a structured way. However, the review also showed that spatial data often needs to be processed before it can be used for pedestrian analysis, especially when crossings and traffic signals are represented inconsistently.

Finally, the chapter discussed urban data platforms and sensing technologies as possible ways to support pedestrian-oriented decision-making. These topics are relevant because the output of this dissertation is intended to be used as a spatial layer, not only as a set of isolated results. The research gaps identified in this chapter motivate the methodology presented next, which combines geospatial data, measured delay values, route simulation and delay-based LOS classification to evaluate pedestrian intersections in Porto.

Chapter 3

Methodology

This chapter describes the approach developed to evaluate pedestrian conditions at urban intersections in Porto. The proposed workflow combines open geospatial data, field measurements, route simulation and delay-based classification to support a spatial assessment of pedestrian delay.

3.1 Introduction

The work is structured around the main stages needed to move from raw spatial data to an intersection-level assessment. Its purpose is to show how the different parts of the process fit together, from the preparation of the spatial data to the calculation of delay-based indicators.

The following sections begin with the study area and the data sources used in the work. The chapter then describes the extraction of intersections and pedestrian crossings, the integration of field observations, the simulation of local pedestrian routes and the classification of the evaluated intersections. The last part explains how the results are exported as spatial data.

3.2 Methodological Overview

Figure 3.1 presents the complete methodological workflow proposed in this dissertation. The workflow is organized as a sequence of processes, from the extraction of spatial data to the production of the final intersection and zone-level outputs.

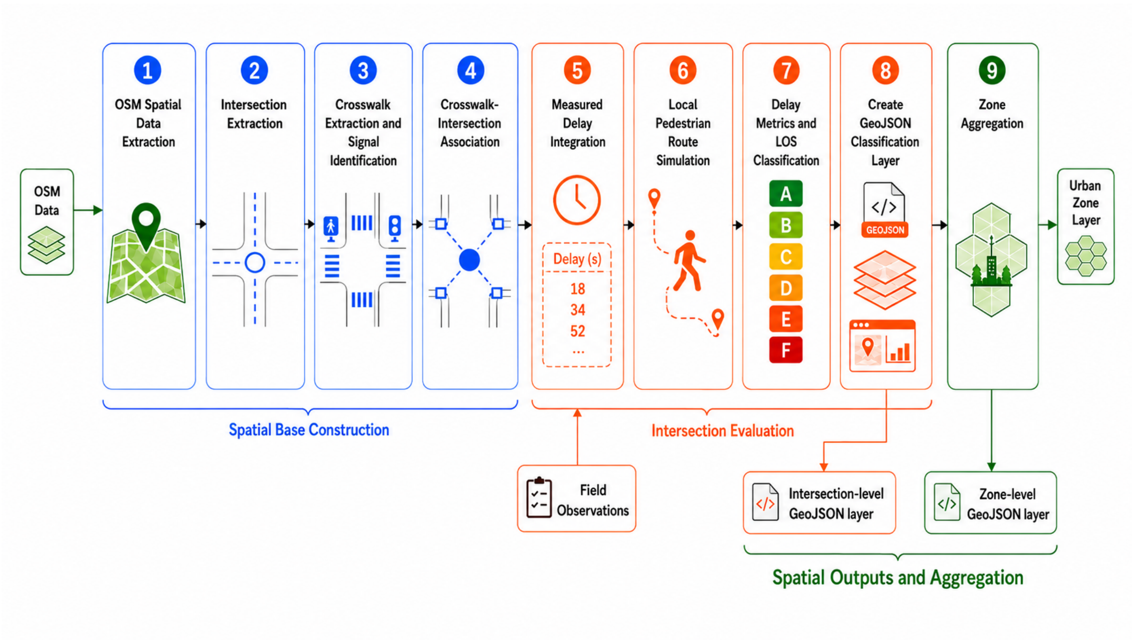

Figure 3.1: Overview of the proposed methodology, showing the main processing steps, input data sources and final spatial outputs. Figure generated with the support of ChatGPT and manually reviewed by the author.

The steps shown in Figure 3.1 can be summarized as follows:

1. **OSM spatial data extraction:** OSM data is used to obtain the road network, pedestrian crossing information and traffic signal information of Porto.
2. **Intersection extraction:** the road network is processed to identify relevant urban intersections, which are used the main spatial units of analysis.
3. **Crosswalk extraction and signal identification:** pedestrian crossings are detected and classified according to the presence or absence of pedestrian signal control.
4. **Crosswalk–intersection association:** detected crossings are linked to nearby intersections, allowing each intersection to store information about its associated pedestrian crossings.
5. **Measured delay integration:** selected field observations are added to the corresponding signalized pedestrian crossings as measured delay values.
6. **Local pedestrian route simulation:** possible pedestrian movements are simulated around each evaluated intersection using the produced pedestrian graph and generated test points.
7. **Delay metrics and LOS classification:** delay indicators are calculated from the simulated routes and translated into pedestrian LOS classes.
8. **GeoJSON export and output layer preparation:** the classified intersections are exported as a spatial layer prepared for visualization and possible integration into the urban monitoring platform used by the city of Porto.

9. **Zone aggregation:** the intersection-level results are associated with urban zones, producing a second spatial output at zone level.

These steps are grouped into three main parts. Steps 1 to 4 build the spatial base of the analysis, by transforming OSM data into a structured set of intersections and pedestrian crossings. Steps 5 to 7 correspond to the evaluation stage, where measured delay values are introduced, pedestrian routes are simulated and intersections are classified. Steps 8 and 9 prepare the final spatial outputs, first at intersection level and then at zone level.

The following sections describe each part of the workflow in more detail, following the order of the steps presented in Figure 3.1.

3.3 Study Area

The study area considered in this dissertation is the municipality of Porto, located in northern Portugal and part of the Porto Metropolitan Area. The Porto Metropolitan Area is one of the main urban regions in Portugal, with a resident population of 1,802,664 inhabitants in 2023 and an area of approximately 2,041 km^2 [45]. Within this metropolitan context, Porto is a compact urban municipality, with an area of approximately 41.42 km^2, a resident population of 231,828 inhabitants in 2021 and a population density of 5,597 inhabitants per km^2 [46]. This density makes pedestrian mobility particularly relevant, since daily movement depends on short-distance connections, access to public transport and frequent interaction between pedestrians and road traffic.

Porto also has urban characteristics that make the assessment of pedestrian intersections especially interesting. The city has an irregular street structure, narrow streets in several areas, frequent intersections and relevant elevation changes. The municipality also presents relevant elevation variation, with available topographic data indicating elevations from approximately sea level to around 210 m [47]. These conditions can affect walking routes, make pedestrian movement less direct and increase the importance of crossings and signalized intersections.

The choice of Porto is also related to the local context of digital urban management. Porto Digital, formally Associação Porto Digital (APD), is a private non-profit association that supports the city's digital transformation and urban innovation initiatives. Its activity is organized around technology, innovation and experimentation and partnerships, with the goal of improving public services and quality of life through digital transition [13]. One example of this context is the Porto Urban Platform, which brings together results from national and European projects related to smart city development [12]. This makes Porto a suitable case for a project that produces geospatial outputs that can later be used in urban monitoring and decision-support tools.

In this work, the analysis is applied to the city network extracted from OSM. From this network, intersections and pedestrian crossings are identified and organized for later evaluation. The delay-based assessment focuses on intersections with signalized pedestrian crossings, since these are the locations where pedestrian waiting time can be associated with traffic signal operation.

In a later stage, the intersections are associated with urban zones provided by the Mobility Department. This allows the results to be read at two spatial levels: individual intersections and the city zones where those intersections are located.

3.4 Data Sources

The workflow uses three main groups of data: open spatial data from OSM, field observations collected at selected crossings and an urban zone layer provided by the city's Mobility Department. Each source is used at a different stage of the process. OSM provides the spatial base, the field observations provide delay information and the zone layer supports the final aggregation of the results.

3.4.1 OpenStreetMap Data

OSM is used as the main source of spatial information in this dissertation. From OSM, the workflow extracts the road network of Porto, including nodes and edges, as well as pedestrian crossings and traffic signal information when available. These elements provide the base from which intersections and crossings are later identified.

The extracted road network is represented as a graph, where nodes and edges describe the structure of the urban network. This representation is important because the following stages depend on identifying where intersections are located and how pedestrian movement can occur around them. Figure 3.2 shows the road network extracted from OSM for the study area.

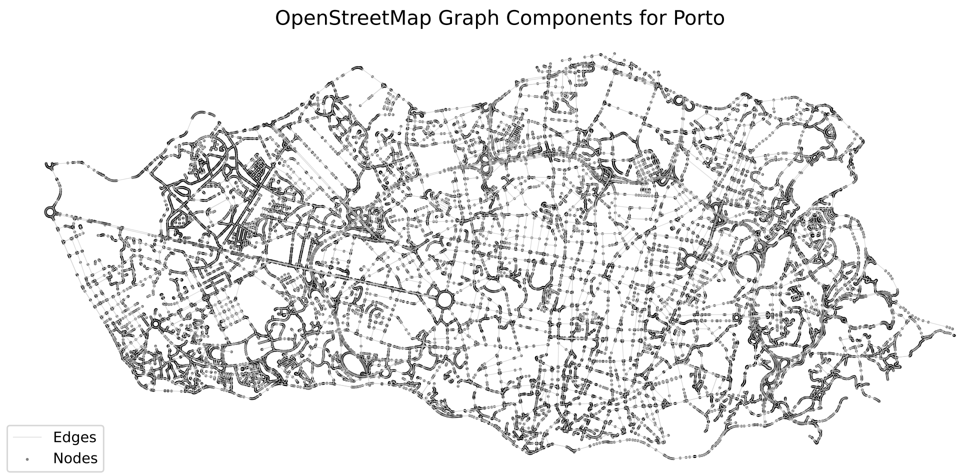

Figure 3.2: Road network extracted from OSM, represented through nodes and edges, for the municipality of Porto. Source: Author

Although OSM provides a useful base, the data is not directly ready for the assessment developed in this work. Intersections, crossings and traffic signals need to be filtered, transformed and connected before they can be used together. This processing is described in the following sections.

3.4.2 Field Observations

The second data source consists of field observations collected at signalized pedestrian crossings. These observations were used to obtain delay-related information associated with specific crossings. The purpose of this data is to add measured waiting time values to the spatial structure extracted from OSM.

The field observations are stored in tabular form and linked to the corresponding crossing and intersection identifiers. This makes it possible to associate measured delay values with the signalized crossings used in the later evaluation. The delay information is therefore not treated as an independent dataset, but as an attribute connected to specific pedestrian crossing locations.

3.4.3 Urban Zone Layer

The third data source is the urban zone layer provided by the Mobility Department of the Porto City Council. This entity is responsible for the mobility-related planning and management within the municipality. From this point onward, it is referred to as the Mobility Department. This layer divides the city of Porto into spatial units that are used in the final aggregation stage of the workflow. Instead of keeping the results only at intersection level, the zone layer allows the classified intersections to be grouped according to the area where they are located.

Figure 3.3 shows the spatial division used in this dissertation. The zones are not generated from the OSM network. They are used as an external spatial reference that supports the interpretation of the intersection-level results at a broader scale.

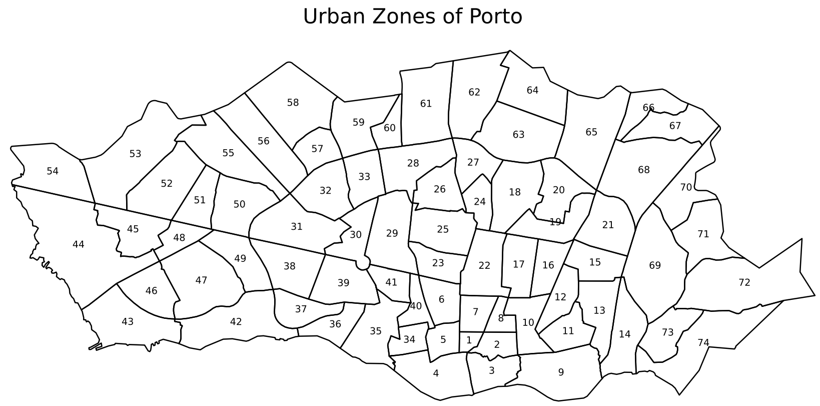

Figure 3.3: Spatial division of Porto into urban zones provided by the Mobility Department. Source: Author

3.4.4 Generated Intermediate Files

Several intermediate files are generated throughout the workflow. These include the extracted nodes and edges, the intersection layer, the crossing datasets, the processed intersection layer and the final zone-level layer. Most of these files are stored as GeoJSON layers or Comma-Separated Values (CSV) tables, depending on whether the information is spatial or tabular.

These files are not external data sources, but they are important for the organization of the workflow. They store the output of each stage and are used as inputs for the following steps. This also makes the process easier to inspect, since each intermediate result can be checked before moving to the next part of the analysis.

Table 3.1 summarizes the main data sources used in the methodology and their role in the workflow.

Table 3.1: Main data sources used in the methodology.

Data source	**Main content**	**Role in the methodology**
OSM	Road network, nodes, edges, pedestrian crossings and traffic signal information	Provides the spatial base used to extract intersections, crossings and network elements.
Field observations	Delay-related information for selected signalized pedestrian crossings	Provides measured delay values used in the evaluation and classification of intersections.
Urban zone layer	Spatial division of the city into mobility zones	Allows intersections to be aggregated and classified at zone level and visualized at higher zoom levels.
Generated intermediate files	Intersections, nodes, edges, crossings with signals, crossings without signals and classified intersections	Store the outputs of each processing step and serve as inputs for the following stages of the workflow.

3.5 Intersection Extraction

As mentioned in previous sections, intersections are used as the main unit of analysis because they represent the places where pedestrian routes are more likely to interact with road traffic, crossings and signal control. Before any delay evaluation can be made, these locations need to be identified in a consistent way.

The extraction starts from the road network obtained from OSM. This network is represented as a graph, composed of nodes and edges, as shown in Figure 3.2. In this structure, nodes can represent points such as street junctions or geometry changes, while edges represent the connections between them. However, the raw graph cannot be used directly as an intersection layer, because not every node corresponds to a relevant data point. Some nodes only exist to preserve the geometry of a street segment, while others may represent very small changes in direction.

Candidate intersections are selected according to the connectivity of each node. In this work, a node is considered a candidate intersection when it is connected to at least three streets. This criterion removes many geometry-only nodes and keeps the focus on locations where different movements meet. Figure 3.4 illustrates this logic. Raw OSM graph nodes are first filtered according to street connectivity, and the retained candidate nodes are then grouped and consolidated into one representative point for each physical intersection.

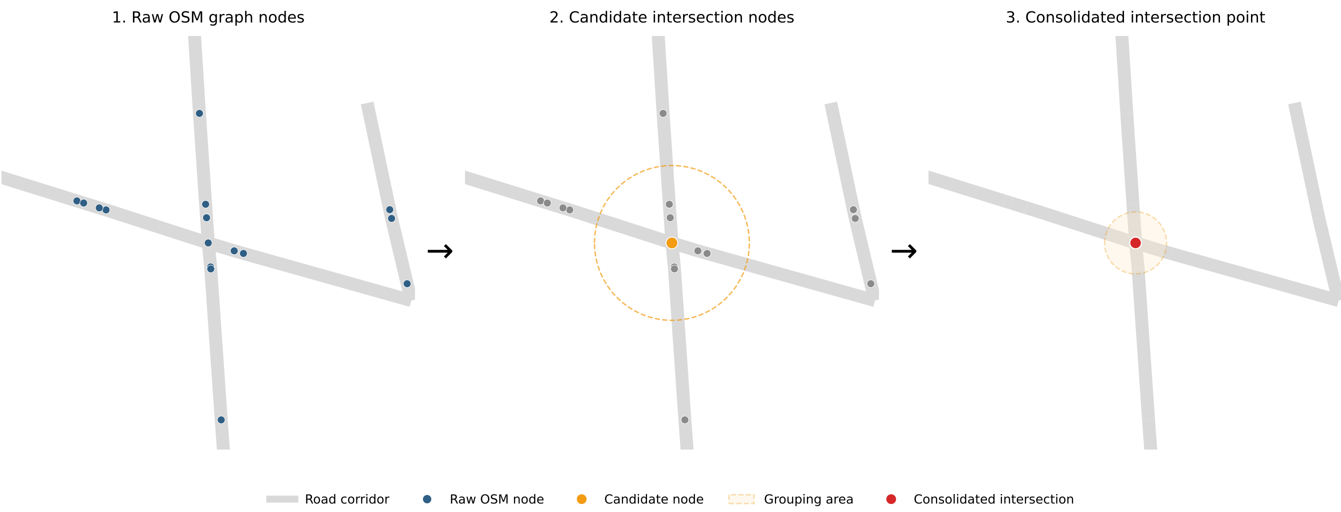

Figure 3.4: Conceptual representation of the intersection extraction process. Raw OSM graph nodes are first filtered according to street connectivity and then consolidated into a single representative point for each physical intersection. Source: Author

The selected nodes are then consolidated to avoid representing the same physical intersection more than once. This is needed because complex road layouts can appear in the graph as several nearby nodes, especially when lanes or street segments are represented separately. Nearby candidate nodes are grouped using a spatial tolerance of 12 m and replaced by a single representative point.

A final filtering step removes almost overlapping intersection points. A minimum distance of 5 m is used between retained intersections, which mainly removes duplicates created during the extraction and consolidation process. All distance-based operations are carried out in a projected coordinate system, so that buffers and proximity checks are calculated in metres.

After this process, each retained intersection is assigned a unique identifier. This identifier is used in the following stages to connect intersections with pedestrian crossings, signal information, delay values, simulated routes and classification results. The output of this step is a spatial layer of extracted intersections, which provides the base for the following parts of the workflow.

To verify that the filtering and consolidation steps did not remove relevant locations, the resulting intersection layer was inspected visually over the road network. This validation was carried out by comparing the consolidated points with the underlying street geometry and by checking whether the retained locations corresponded to real junctions in the city network. Particular attention was given to dense areas, where the raw OSM graph contains many nearby nodes and where over-cleaning could remove valid intersections. This inspection showed that the consolidation process reduced duplicated graph nodes while preserving the main physical intersections needed for the following stages.

It is important to note that this validation was based on the available OSM representation. Therefore, the quality of the result depends on the completeness and accuracy of the mapped street network. However, the manual inspection provided confidence that the adopted thresholds produced a coherent intersection layer for the purposes of this dissertation.

3.6 Crosswalk Extraction, Signal Identification and Association

After extracting the intersections, the next step is to identify the pedestrian crossings associated with them. This step is necessary because the delay evaluation developed in this dissertation depends on the existence of signalized pedestrian crossings. Intersections without pedestrian crossings, or without crossings controlled by pedestrian signals, are not the main focus of the delay-based classification.

Pedestrian crossings are extracted from OSM using crossing-related tags. In the original data, crossings may be represented in different ways. Some appear as points, while others are represented as line geometries or more complex geometries. To make the data easier to use in the following stages, each crossing is converted into a representative point. This point is used to associate the crossing with nearby intersections and later, with delay information.

Each crossing is assigned a unique crossing identifier. This is important because the next steps need to connect several types of information: the crossing geometry, the associated intersection, the presence of pedestrian signal control and the delay value when available. Without this identifier, it would be harder to track the same crossing across the different files generated by the workflow.

The identification of signalized pedestrian crossings is based on two types of information. First, the method checks crossing-related tags in OSM, such as attributes indicating the presence of pedestrian signals. Second, it checks the proximity between crossings and traffic signal features. This proximity step is needed because signal information is not always mapped directly on the crossing feature itself.

After crossings are prepared, they are associated with nearby intersections. A spatial search is used around each extracted intersection to identify crossings located in its surroundings. This association step links each crossing to the intersection where it is likely to affect pedestrian movement.

A deduplication step is also applied to avoid counting the same crossing more than once. This is needed because the same physical crossing may appear in OSM through more than one geometry or through very close mapped elements. When duplicated crossings are found, the retained crossing keeps the relevant signal information if any of the duplicated elements was identified as signalized.

The output of this step is a set of crossings associated with intersections and classified according to the presence or absence of pedestrian signal control. This information is then used to select the intersections where pedestrian delay can be evaluated.

This also means that the workflow becomes progressively more selective. A location may be a valid street intersection in the road network, but it only continues to the delay evaluation stage if it is associated with at least one pedestrian crossing and, later, with pedestrian signal control.

For example, a junction between several streets may be retained during the intersection extraction stage, but if no pedestrian crossing is identified nearby in the OSM data, it is excluded from the signalized pedestrian delay assessment. In this case, the location is not considered wrong, it is simply outside the scope of the final evaluation, which focuses on signalized pedestrian crossings.

3.7 Measuring Pedestrian Delay

After identifying the signalized pedestrian crossings, delay values are added to the few crossings where field observations were possible to collect. In this dissertation, pedestrian delay is treated as a crossing-level attribute, since waiting time occurs when a pedestrian reaches a signalized crossing and waits for permission to cross.

The measured values are linked to the crossing identifiers created in the previous stage. This connection allows each observed delay to remain associated with a specific crossing and, through the crossing–intersection association, with the corresponding intersection.

This structure is important because an intersection may contain more than one signalized pedestrian crossing. Keeping the delay at crossing level allows the following route simulation stage to assign different waiting times to different pedestrian movements around the same intersection.

The output of this step is a set of signalized pedestrian crossings with associated delay values where observations are available. These values are then used in the route simulation stage that is described in the next section.

3.8 Pedestrian Route Simulation

After the intersections with signalized pedestrian crossings are identified and the delay information is associated with them, the next step is to evaluate how these delays can affect pedestrian movement around each intersection. For this purpose, the methodology uses a controlled pedestrian route simulation.

The simulation does not aim to reproduce the exact behaviour of every pedestrian. Instead, it is used to test possible movements around an intersection and to estimate how signalized crossings can influence those movements. This distinction is important because the objective is not to model individual pedestrian decisions, but to create a reproducible way of connecting crossing-level delay values with intersection-level indicators.

The route simulation is applied at the intersection level. Around each intersection, a set of test points is generated to represent possible pedestrian origins and destinations in the surrounding area. These points do not represent observed pedestrians. They are reference points used to create possible crossing movements around the intersection.

Routes are then simulated between pairs of test points using the prepared pedestrian network. This allows the workflow to identify which pedestrian crossings are used in each path when moving through the intersection. When a simulated route uses a signalized crossing with associated delay information, that delay is added to the route as a time penalty.

This step is useful because the delay is first associated with individual crossings, while the assessment is produced at intersection level. A single intersection may include several crossings, and different pedestrian movements may be affected by different waiting times. Route simulation therefore helps connect crossing-level delays to the broader operation of the intersection.

The output of this step is a set of simulated routes and associated delay values for each intersection. These values are then used to calculate the delay-based indicators described in the following section.

3.9 Delay Metrics and Level of Service Classification

After the pedestrian routes are simulated, the delay values obtained for each route are used to calculate intersection-level indicators. This step converts the crossing-level delay information into a classification that can be assigned to each evaluated intersection.

Two main delay indicators are calculated for each intersection. The first is the average delay, which represents the mean delay across the simulated routes associated with that intersection. This value gives a general indication of the conditions experienced in the pedestrian movements. The second is the maximum delay, which represents the highest delay found among the simulated routes. This value is useful because as it is known, a single intersection may contain several signalized crossings and some pedestrian movements may be more affected than others.

The use of both indicators helps describe the intersection in a more balanced way. The average delay provides an overall reading of the intersection, while the maximum delay identifies the most critical simulated movement. This distinction is important because an intersection may have acceptable average conditions while still containing one crossing or route with a critical waiting time.

Based on these delay indicators, each evaluated intersection is classified according to pedestrian LOS. In this dissertation, LOS is used as a simplified qualitative scale to translate delay values into interpretable classes, ranging from LOS A to LOS F. Lower delay values correspond to better pedestrian conditions, while higher delay values correspond to worse conditions.

The classification thresholds used in this work are shown in Table 3.2. These thresholds were defined as part of the proposed methodology and are not intended to reproduce a specific standard or official LOS scale. Instead, they provide a consistent rule for comparing the evaluated intersections within this case study and for converting delay values into classes that can be visualized and aggregated in the urban monitoring platform. The adopted ranges follow the general assumption that pedestrian delay becomes progressively more disruptive as waiting time increases.

Table 3.2: Pedestrian LOS classification thresholds adopted in this dissertation.

LOS class	**Delay range**	**Interpretation**
A	≤ 20 s	Very low delay. Pedestrian movement is only slightly affected.
B	> 20–40 s	Low delay. Waiting time is still generally acceptable.
C	> 40–60 s	Moderate delay. Pedestrian movement begins to be noticeably interrupted.
D	> 60–90 s	High delay. Waiting time has a clear effect on route continuity.
E	> 90–120 s	Very high delay. Pedestrian movement is strongly interrupted.
F	> 120 s	Excessive delay. Waiting time represents poor pedestrian conditions.

In addition to the LOS class, each intersection is assigned a numerical severity score. This score is derived from the LOS class and allows the classification to be used more easily in later processing steps, such as visualization and zone aggregation. Lower scores represent better conditions, while higher scores represent more severe delay conditions.

The output of this step is a layer where each intersection has an assigned delay indicator, the corresponding LOS class, the severity score and visualization attributes. These attributes are then used in the spatial export stage and in the aggregation of results by urban zone.

3.10 GeoJSON Export and Urban Platform Layer Preparation

After the delay metrics and LOS classification are calculated, the results are exported as a spatial layer. This step is necessary because the assessment is not only a table of values. Each classified intersection has a geographic location and must remain connected to the attributes calculated in the previous stages.

The intersection-level output is exported as a GeoJSON layer in which each feature corresponds to one evaluated intersection. At this stage, the purpose is to preserve the link between the spatial position of the intersection and the attributes calculated in the previous steps. For each point, the exported features include the intersection identifier, average delay, maximum delay, number of simulated routes, number of routes with alert conditions, LOS class, severity score and visualization fields.

The visualization fields are included to make the layer easier to inspect in mapping environments. Each LOS class is associated with a colour, allowing intersections with better or worse delay conditions to be quickly distinguished on a map. This makes the output more useful for spatial interpretation, since the classification can be read visually without needing to inspect each attribute individually.

The GeoJSON structure also supports a smooth integration with the urban monitoring platform used by the city of Porto. For this reason, the exported layer keeps the information in a format that can be connected to web-based geospatial visualization tools. The goal is not only to produce a final analytical result, but also to organize the result as a practical urban data layer.

This export stage also helps separate the methodological process from the final product. The previous stages generate and classify the information, while this stage prepares the classified intersections for visualization, inspection and future use in urban monitoring contexts.

3.11 Zone Aggregation

After the intersection-level classification is produced, the results are aggregated by urban zones. This step adds a broader spatial reading to the assessment, allowing the results to be interpreted not only as individual points, but also as part of larger areas within the city.

The aggregation uses the urban zone layer described in the data sources section and shown in Figure 3.3. Each intersection from the classification layer is associated with the zone where it is located through a spatial join. This creates a connection between the point-based intersection layer and the polygon-based zone layer.

However, not all intersections in the intersection-level layer are used to calculate the zone indicators. Currently, the intersection layer also includes points without associated delay values, which are kept in the output to show where the method can be applied in the future. These intersections do not have a valid LOS classification and therefore should not influence the zone-level results. For this reason, the zone aggregation only considers intersections with a valid LOS class.

This distinction is important because missing delay information may appear as a default numerical value in some fields. A delay value equal to zero is therefore not enough to indicate good pedestrian conditions. In this methodology, an intersection is only considered valid for aggregation when it has a valid LOS class. This prevents intersections without real delay measurements from being interpreted as low-delay results.

Once the valid intersections are linked to zones, a numerical score is assigned to each one according to its LOS class. The score ranges from 1 to 6, where 1 corresponds to LOS A and 6 corresponds to LOS F. The zone score is then calculated from the average severity score of the valid intersections located inside each zone.

In addition to the average score, other summary indicators are retained. These include the worst score, the best score, the number of valid intersections used in the aggregation, the zone-level LOS and the worst LOS found inside the zone. These values help interpret the aggregated result, since a zone with only one valid intersection should not be read in the same way as a zone with several valid intersections.

Zones without any valid LOS-classified intersection are treated as zones without data. They are kept in the output layer for visualization and spatial continuity, but they do not receive an analytical classification. This makes the absence of measured information explicit instead of assigning an artificial good classification to those areas.

The output of this step is a zone-level spatial layer. Each feature represents one urban zone and stores the aggregated pedestrian delay indicators associated with that area. This feature does not replace the intersection-level assessment. Instead, it provides a second scale of analysis. The intersection layer keeps the detailed local information, while the zone layer supports a more general reading of pedestrian delay conditions across the city.

3.12 Final remarks

This chapter described the methodology used to transform open spatial data and field observations into a pedestrian delay assessment for urban intersections in Porto. The workflow starts with the extraction and preparation of the road network, intersections, pedestrian crossings and traffic signal information. These steps create the spatial base needed to identify the locations where pedestrian delay can be evaluated.

The methodology then focuses on intersections with signalized pedestrian crossings, since these are the places where pedestrian movement is controlled by traffic signals and where waiting time can be introduced. Delay information is associated with the corresponding crossings and used in a local pedestrian route simulation. This allows delay values to be related to pedestrian movements around each intersection.

From the simulated routes, delay-based indicators are calculated and used to assign a pedestrian LOS classification to each intersection. The results are then exported as an intersection-level GeoJSON layer, prepared for spatial visualization and possible future integration into the urban monitoring platform used by the city of Porto.

Finally, the methodology includes an aggregation step using the urban zone layer. This makes it possible to read the results at two spatial scales: the detailed intersection level and the broader city-zone level. Together, these outputs provide a structured way to represent delay conditions and create a base that can be updated in the future as more measured or sensor-based data becomes available.

Chapter 4

Implementation

This chapter describes how the methodology presented in Chapter 3 was implemented. The implementation was organized as a sequence of computational steps that follow the workflow described in the previous chapter. Each step produces intermediate outputs that are used by the following stage, allowing the process to move from raw spatial data to classified intersection and zone layers.

The first part of the implementation extracts and prepares the spatial base of the study. This includes the road network, intersections, pedestrian crossings and signalized crossing information. The second part uses this prepared data to simulate local pedestrian routes, assign delay values, calculate delay indicators and classify intersections according to pedestrian LOS. The final part aggregates the classified intersections by urban zones and prepares the spatial outputs for visualization and possible integration into the urban monitoring platform used by the city of Porto.

Therefore, this chapter presents the implementation from a technical perspective. It describes the used tools, the adopted coordinate reference systems, the main parameters defined in the processing workflow and the structure of the files generated throughout the work.

4.1 Development Environment and Tools

The implementation of the proposed approach was developed in Python, using a set of libraries suited for geospatial data processing, network analysis and tabular data manipulation. The main libraries used in the workflow are summarized in Table 4.1.

Table 4.1: Main Python libraries used in the implementation.

Library	**Main purpose**	**Use in this dissertation**
OSMnx	Extraction and processing of OSM network data	Downloading and preparing the road network used as the spatial base for intersection extraction and route simulation.
GeoPandas	Geospatial data processing	Reading, transforming, filtering, joining and exporting spatial layers such as intersections, crossings, nodes, edges and zones.
Shapely	Geometric operations	Creation and manipulation of geometries, including points, lines, buffers, representative points, nearest points and spatial unions.
Pandas	Tabular data processing	Management of crossing tables, delay values, identifiers, intermediate results and summary outputs.
NetworkX	Graph-based analysis	Representation and analysis of network structures used during pedestrian route simulation.
NumPy	Numerical array operations	Support for numerical calculations involving lists or arrays of values, such as distance, delay and route-related computations.
Math	Basic mathematical functions	Support for scalar mathematical operations, including trigonometric and angle-related calculations.
Itertools	Combination generation	Creation of pairs or combinations of points used in the route simulation stage.
Pathlib and os	File and path management	Definition of input and output paths and organization of generated files across the project folders.

Coordinate reference systems were also an important part of the implementation. The data extracted from OSM is originally represented in European Petroleum Survey Group (EPSG):4326, a geographic coordinate reference system based on World Geodetic System 1984 (WGS84). This system is widely used as a de facto standard for geographic data exchange, web mapping and formats such as GeoJSON [17]. However, distance-based operations cannot be carried out reliably in this coordinate system because the units are expressed in degrees.

For this reason, all metric operations were performed in EPSG:3763, the projected coordinate reference system used for mainland Portugal. This includes operations such as buffers, proximity searches, distance thresholds and deduplication of nearby geometries. After processing, the final spatial layers were converted back to EPSG:4326 so that they could be exported and visualized in web mapping environments.

The implementation was also structured to generate intermediate outputs after each major processing step. This made the workflow easier to inspect and validate, since intersections, crossings, network elements and final classifications could be checked separately before being used in the next stage.

4.2 Implementation Workflow

The implementation was organized into three main processing stages, following the methodology described in Chapter 3. The first stage extracts and prepares the spatial data required for the analysis. The second stage uses this spatial base to simulate pedestrian routes, calculate delay indicators and classify intersections. The third stage aggregates the classified intersections by urban zones.

This structure makes the workflow easier to inspect and validate, since each stage produces intermediate outputs that are used in the following stage. It also separates the main technical tasks of the implementation: spatial data extraction, delay-based intersection classification and zone-level aggregation.

4.2.1 Spatial Data Extraction

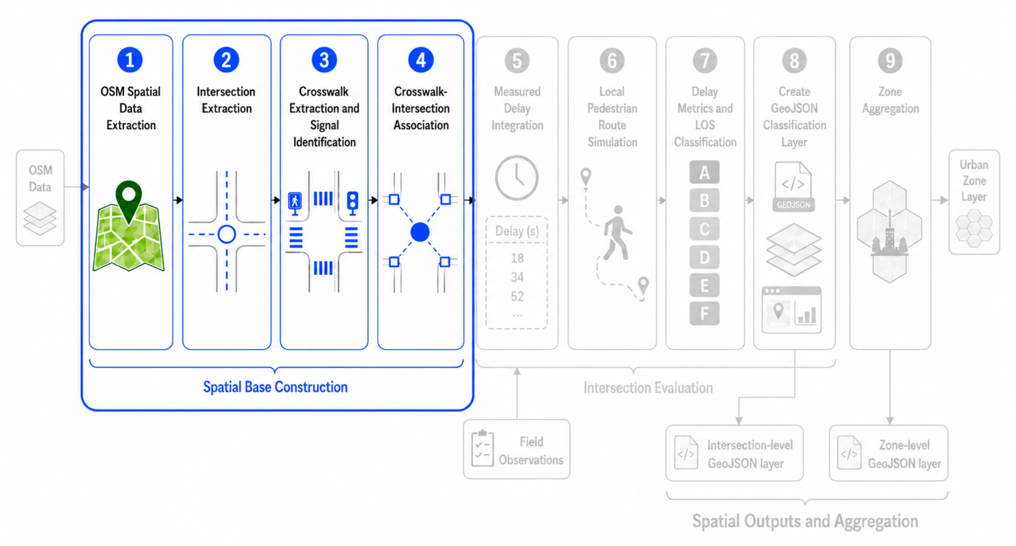

Figure 4.1: Workflow phase corresponding to spatial base construction. Figure generated with the support of ChatGPT and manually reviewed by the author.

The spatial data extraction stage implements the first part of the workflow described in Chapter 3 and highlighted in Figure 4.1. Its purpose is to transform OSM data into a set of structured spatial layers that can be used by the simulation and classification stages.

The parameters used in this stage are presented before the detailed explanation of the workflow, since they define the thresholds used for intersection filtering, node consolidation, crossing association, signal proximity checks and deduplication. These values are summarized in Table 4.2.

Table 4.2: Main parameters used in the spatial data extraction stage.

Parameter	**Value**	**Purpose**
`MIN_STREETS`	3	Minimum number of connected streets required for a graph node to be considered a candidate intersection.
`TOLERANCE`	12 m	Distance used to consolidate nearby candidate nodes that may belong to the same physical intersection.
`MIN_DIST`	5 m	Minimum distance allowed between retained intersection points after consolidation.
`CROSSWALK_SEARCH_RADIUS`	20 m	Search distance used to associate pedestrian crossings with nearby intersections.
`PED_SIGNAL_PROXIMITY_RADIUS`	12 m	Distance used to associate crossings with nearby traffic signal features when signal information is not stored directly on the crossing.
`CROSSWALK_DEDUP_DIST`	6 m	Distance used to identify duplicated or nearly duplicated crossings associated with the same intersection.

The road network is extracted with OSMnx and converted into GeoDataFrames of nodes and edges, as shown in Algorithm 1, lines 1–3. The graph is first obtained in EPSG:4326 and then projected to EPSG:3763, since the operations in this stage depend on metric distances. These operations include intersection consolidation, crosswalk-intersection association, traffic signal proximity checks and deduplication. After processing, the exported spatial layers are converted back to EPSG:4326.

The implementation uses the graph node attribute related to street connectivity to identify candidate intersections. Nodes with a street count lower than `MIN_STREETS` are discarded, as represented in Algorithm 1, lines 4–9. The retained candidates are then grouped according to the consolidation tolerance and each group is converted into one representative intersection point. A second distance filter is applied to remove almost overlapping points that remain after consolidation, as shown in lines 10–18. Figure 4.2 illustrates this consolidation step. Several raw OSM nodes located around the same physical junction are grouped and represented by a single consolidated intersection point.

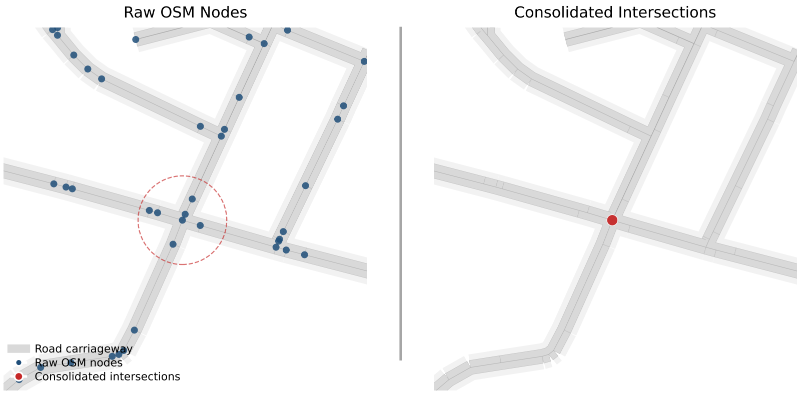

Figure 4.2: Consolidation of raw OSM nodes into a single representative intersection point. Source: Author

Pedestrian crossings are extracted separately from OSM using the tags `highway=crossing` and `footway=crossing`. Because these features may be represented as points, lines, multilines, or other geometry types, each crossing is converted into a representative point before being used in spatial operations. Each crossing also receives a unique `cross_id`. When possible, this identifier preserves the original OSM element type and identifier, which avoids confusion between nodes, ways and relations that may share the same numeric ID. These steps correspond to Algorithm 1, lines 19–24.

Signalized crossings are detected through two complementary procedures. The first checks the attributes of the crossing itself, namely `crossing`, `crossing:signals` and `traffic_signals`. A crossing is classified as signalized when one of these fields contains values such as `traffic_signals`, `pedestrian_signals`, `yes`, `true`, or `1`. The second procedure checks whether the crossing point is located near a mapped `highway=traffic_signals` feature. This is needed because signal information is not always stored directly in the crossing feature. The combination of tag-based and proximity-based detection is summarized in Algorithm 1, lines 25–28.

The association between crossings and intersections is carried out with a local radius-based search, as shown in Algorithm 1, lines 30–36. For each intersection, only crossings located inside the defined search radius are considered. This avoids using a global nearest-neighbour rule, which could incorrectly assign a crossing to the closest intersection even when it is outside the local intersection area. Figure 4.3 illustrates the local association logic. Only crossings located within the defined search area around the consolidated intersection are associated with that intersection.

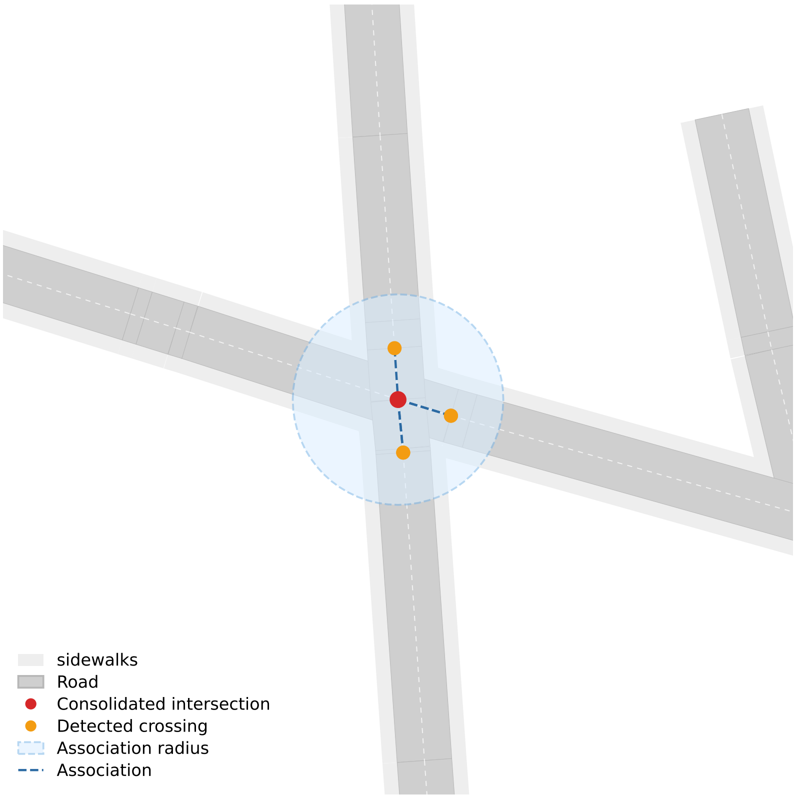

Figure 4.3: Association between a consolidated intersection and nearby pedestrian crossings using a local search radius. Source: Author

Finally, crossings associated with the same intersection are deduplicated, as represented in Algorithm 1, line 32. Crossings closer than the defined deduplication distance are treated as possible representations of the same physical crossing. If at least one element in the duplicated group is signalized, the retained crossing keeps the signalized status. This prevents signal information from being lost during the cleaning process. The final signalized and non-signalized crossing datasets are then separated and exported, as shown in lines 37–39.

The main logic of this stage is summarized in Algorithm 1. The pseudocode abstracts the implementation details while preserving the main operations and decision rules used to transform OSM data into the spatial layers required by the following stages.

Algorithm 1 Spatial data extraction and crossing association

Require: OSM road network, crossing features and traffic signal features for Porto
Require: `MIN_STREETS`, `TOLERANCE`, `MIN_DIST`, `CROSSWALK_SEARCH_RADIUS`, `PED_SIGNAL_PROXIMITY_RADIUS`, `CROSSWALK_DEDUP_DIST`
Ensure: Intersection layer, network layers, signalized crossing dataset and non-signalized crossing dataset

$G \leftarrow$ extract road network from OSM
$G_{proj} \leftarrow$ project G to EPSG:3763
$(N, E) \leftarrow$ convert G_{proj} into node and edge spatial layers
$C \leftarrow \emptyset$
for all $n \in N$ **do**
  if $street\_count(n) \geq$ `MIN_STREETS` **then**
    $C \leftarrow C \cup \{n\}$
  end if
end for
$K \leftarrow$ cluster nodes in C using `TOLERANCE`
$I \leftarrow \emptyset$
for all $k \in K$ **do**
  $p \leftarrow$ representative point of cluster k
  if $distance(p, I) \geq$ `MIN_DIST` **then**
    assign unique `intersection_id` to p
    $I \leftarrow I \cup \{p\}$
  end if
end for
$W \leftarrow$ extract pedestrian crossings from OSM
$S \leftarrow$ extract traffic signal features from OSM
$P \leftarrow \emptyset$
for all $w \in W$ **do**
  $q \leftarrow$ representative point of crossing w
  assign unique `cross_id` to q
  $tagSignal \leftarrow$ check crossing tags of w for pedestrian signal information
  $nearSignal \leftarrow \exists s \in S : distance(q, s) \leq$ `PED_SIGNAL_PROXIMITY_RADIUS`
  $q.has\_signal \leftarrow tagSignal \lor nearSignal$
  $P \leftarrow P \cup \{q\}$
end for
for all $i \in I$ **do**
  $P_i \leftarrow \{q \in P : distance(q, i) \leq$ `CROSSWALK_SEARCH_RADIUS`$\}$
  $P_i^{dedup} \leftarrow$ deduplicate P_i using `CROSSWALK_DEDUP_DIST`
  for all $q \in P_i^{dedup}$ **do**
    assign $i.intersection\_id$ to q
  end for
end for
$P_{signalized} \leftarrow \{q \in P : q.has\_signal = true\}$
$P_{nonSignalized} \leftarrow \{q \in P : q.has\_signal = false\}$
export I, N, E, $P_{signalized}$ and $P_{nonSignalized}$

The outputs generated in this stage are summarized in Table 4.3. These files are used to validate the extraction process and provide the input data for route simulation and delay-based classification.

Table 4.3: Outputs generated during the spatial data extraction stage.

Output	**Format**	**Content**
`nodes.geojson`	GeoJSON	Road network nodes extracted from OSM.
`edges.geojson`	GeoJSON	Road network edges extracted from OSM.
`intersections.geojson`	GeoJSON	Consolidated intersection points used as the spatial base for the assessment.
`crosswalks_with_lights.csv`	CSV	Signalized pedestrian crossings with identifiers and an editable delay field.
`crosswalks_without_lights.csv`	CSV	Non-signalized pedestrian crossings kept for inspection and context.
`crosswalks_with_lights.geojson`	GeoJSON	Spatial layer of pedestrian crossings identified as signalized.
`crosswalks_without_lights.geojson`	GeoJSON	Spatial layer of pedestrian crossings identified as non-signalized.

Figures 4.4 and 4.5 show examples of the spatial outputs generated in this stage. Figure 4.4 presents the extracted intersection layer over the OSM road network, while Figure 4.5 shows the extracted pedestrian crossings, distinguishing crossings with and without identified pedestrian signal control. The road network nodes and edges are not repeated as a separate figure in this chapter, since they were already shown in Chapter 3, Figure 3.2.

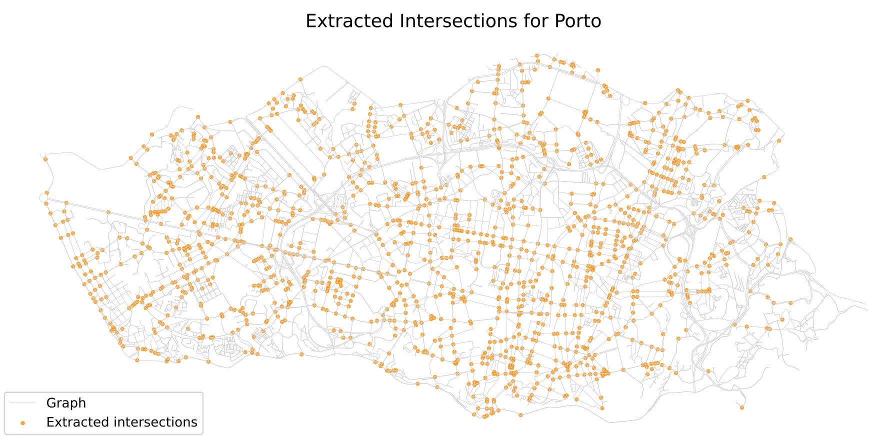

Figure 4.4: Extracted intersections generated during the spatial data extraction stage. Source: Author

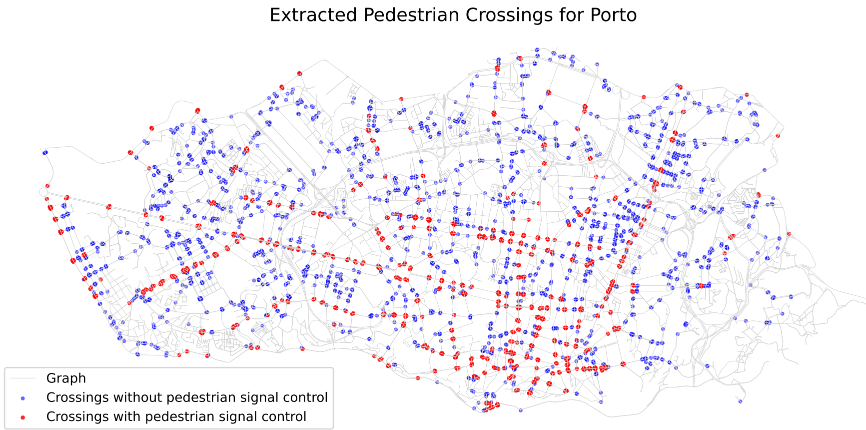

Figure 4.5: Extracted pedestrian crossings, distinguishing crossings with and without pedestrian signal control. Source: Author

4.2.2 Route Simulation and Intersection Classification

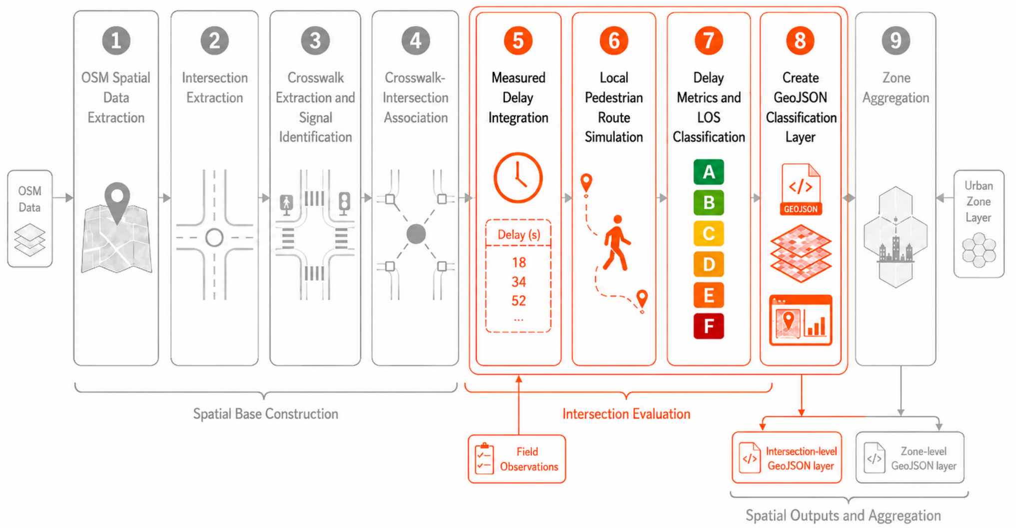

Figure 4.6: Workflow phase corresponding to intersection evaluation. Figure generated with the support of ChatGPT and manually reviewed by the author.

The second processing stage, highlighted in Figure 4.6, uses the spatial outputs generated in the previous stage to simulate local pedestrian movements around the intersections with signalized pedestrian crossings. The goal of this stage is to connect crossing-level delay values to possible pedestrian routes and to produce an intersection-level classification.

The input data for this stage includes the extracted intersections, the road network nodes and edges, the signalized crossing layer and the table containing the measured delay values. The intersection layer is first filtered so that only intersections associated with at least one signalized pedestrian crossing are processed, as shown in Algorithm 2, lines 1–3. This keeps the implementation focused on the locations where pedestrian delay can be related to traffic signal operation. Figure 4.7 shows the subset of extracted intersections that are processed in this stage. These are the intersections associated with at least one signalized pedestrian crossing and therefore considered for delay-based evaluation.

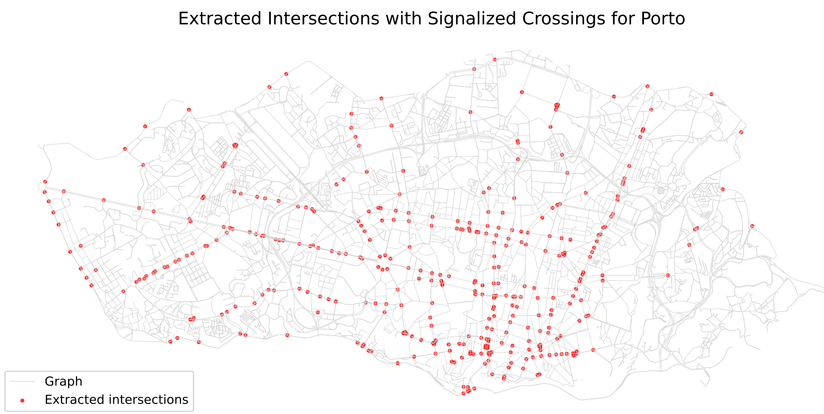

Figure 4.7: Intersections selected for delay-based evaluation, corresponding to the subset of extracted intersections associated with at least one signalized pedestrian crossing. Source: Author

For each intersection, the implementation creates a local working area around the intersection centre, as represented in Algorithm 2, lines 6–9. The road network elements inside this area are extracted from the preloaded network and used to build a local pedestrian graph. This local approach reduces the amount of data processed for each intersection and keeps the route simulation focused on the pedestrian movements that are most relevant to that location.

The pedestrian graph is one of the most important elements of this stage. Rather than representing the intersection only as a central point with nearby crossings, the implementation builds a local pedestrian-oriented graph that preserves walkable paths, crossing locations and the spatial arrangement of the surrounding street layout. This construction is summarized in Algorithm 2, lines 14–15. This makes it possible to simulate pedestrian movement more realistically, including the side of the street used by the pedestrian and the specific crossings involved in each route.

Automatic test points are then generated around each evaluated intersection to represent possible local pedestrian origins and destinations. These points are placed on the pedestrian graph and grouped into intersection arms according to their angular position relative to the intersection centre. Pairs of test points are then used to define candidate pedestrian movements through the intersection area. This process corresponds to Algorithm 2, lines 16–17.

Figure 4.8 illustrates an example of this structure. It shows the local pedestrian graph, the generated test points and one example route calculated over the graph. This helps show how the implementation moves from a local spatial representation of the intersection to route-based pedestrian delay calculations.

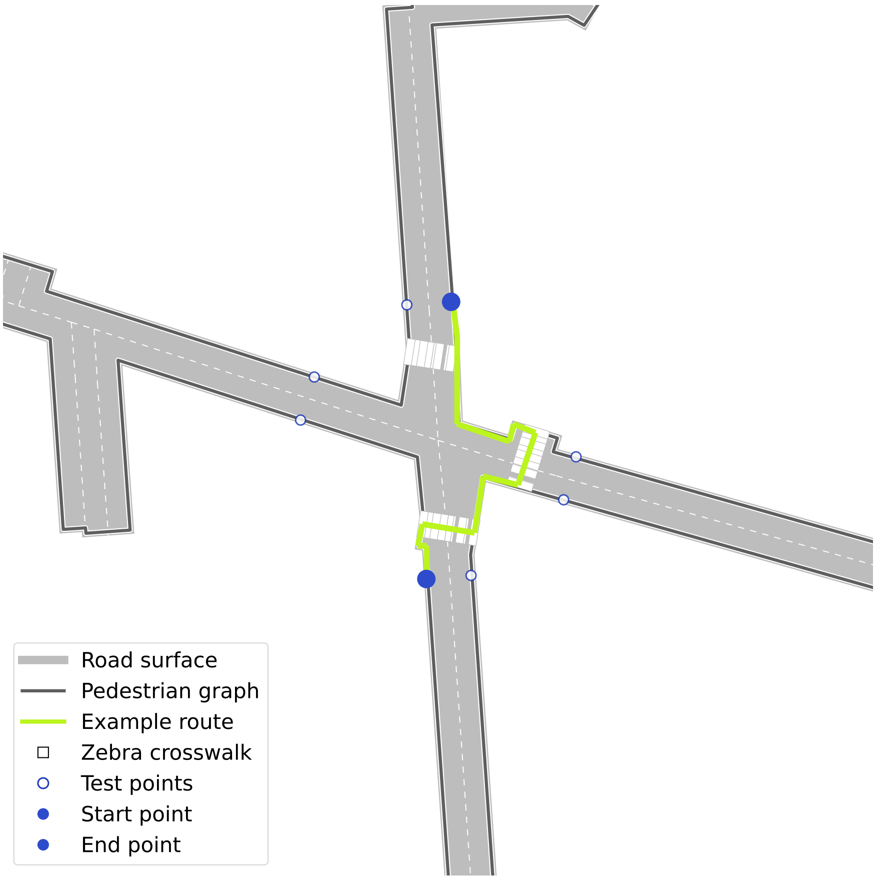

Figure 4.8: Example of a local pedestrian graph constructed around an evaluated intersection, including generated test points and an example route used in the simulation process. Source: Author

For each pair of test points, the implementation attempts to calculate a route through the local pedestrian graph, as shown in Algorithm 2, lines 18–26. If a valid route is found, the route length, walking time, used crossings and accumulated crossing delay are stored. As shown in Figure 4.8, the route follows the pedestrian graph instead of connecting the points with a straight line. This ensures that the simulated movement respects the available walkable paths and crossing locations around the intersection.

The crossing delay is then matched to the official signalized crossing points using a spatial matching distance. This makes it possible to identify which measured crossing delays are actually used by each simulated pedestrian route. Different pairs of test points can therefore produce different delay values, depending on the crossings included in each route. This matching and delay calculation step is represented in Algorithm 2, lines 22–24.

The intersection-level indicators are calculated from the valid simulated routes, as detailed in Algorithm 2, lines 27–30. The mean delay represents the average delay across the routes, while the maximum delay represents the highest delay found among them. The number of tested routes and the number of routes exceeding the alert threshold are also stored.

The LOS classification is assigned from the calculated delay values. Intersections with available delay values receive a LOS class, a corresponding colour, a numerical rank and a severity score. Intersections without real delay values are exported without a LOS class and are assigned a neutral grey colour for visualization. This allows the final layer to keep the full structure of signalized intersections while distinguishing the locations where measured delay values are available.

At the end of this stage, the route-based indicators and the LOS classification are written into the final intersection-level classification layer. Each evaluated intersection is exported as a point feature containing its identifier, delay metrics, number of simulated routes, number of routes with alert conditions, LOS class, severity score, alert status and visualization attributes. This feature creation and export process is shown in Algorithm 2, lines 31–37. It produces the `intersections_evaluated.geojson` file, which is the main output of the intersection evaluation stage and the input used in the following zone aggregation stage.

The main logic of this stage is summarized in Algorithm 2.

Algorithm 2 Route simulation and intersection classification

Require: Intersection layer, network nodes and edges, signalized crossings and delay table
Require: `DOWNLOAD_DIST_M`, `WALK_SPEED_MPS`, `ROUTE_CROSSWALK_MATCH_DIST_M`, `ALERT_DELAY_THRESHOLD_S`
Ensure: Evaluated intersection layer and test point debug layer

$C \leftarrow$ load signalized crossings with delay values
$I \leftarrow$ load extracted intersections
$I_s \leftarrow$ filter I to intersections with signalized crossings in C
$F \leftarrow \emptyset$
$T \leftarrow \emptyset$
for all $i \in I_s$ **do**
  $C_i \leftarrow$ signalized crossings associated with intersection i
  $D_i \leftarrow$ load delay values for crossings in C_i
  $(N_i, E_i) \leftarrow$ crop local network around i using `DOWNLOAD_DIST_M`
  if E_i is empty **then**
    mark i as problematic
    continue
  end if
  $(E_i^{walk}, W_i) \leftarrow$ separate walkable edges and crossing geometries
  $G_i \leftarrow$ build local pedestrian graph from E_i^{walk} and W_i
  $P_i \leftarrow$ generate automatic test points around i
  $R_i \leftarrow$ create route pairs from P_i
  $Routes_i \leftarrow \emptyset$
  for all $(p_a, p_b) \in R_i$ **do**
    $r \leftarrow$ shortest path between p_a and p_b in G_i
    if r is valid **then**
      $usedCrossings \leftarrow$ match route r to crossings in C_i
      $delay_r \leftarrow$ sum delay values of $usedCrossings$
      store route length, route time and $delay_r$ in $Routes_i$
    end if
  end for
  $meanDelay_i \leftarrow$ mean delay of valid routes in $Routes_i$
  $maxDelay_i \leftarrow$ maximum delay of valid routes in $Routes_i$
  $alert_i \leftarrow$ count routes where delay exceeds `ALERT_DELAY_THRESHOLD_S`
  $los_i \leftarrow$ classify $meanDelay_i$ according to LOS thresholds
  $feature_i \leftarrow$ create evaluated intersection feature
  $F \leftarrow F \cup \{feature_i\}$
  $T \leftarrow T \cup P_i$
end for
assign visualization attributes to all features in F
export F as `intersections_evaluated.geojson`
export T as `intersections_test_points.geojson`

The main parameters used in this stage are summarized in Table 4.4.

Table 4.4: Main parameters used in the route simulation and classification stage.

Parameter	**Value**	**Purpose**
DOWNLOAD_DIST_M	70 m	Distance used to crop the local network around each intersection.
WALK_SPEED_MPS	1.2 m/s	Walking speed used to convert route distance into walking time.
ROUTE_CROSSWALK_MATCH_DIST_M	8 m	Distance used to match simulated routes to official signalized crossing points.
DEFAULT_PED_SIGNAL_DELAY_S	0 s	Default delay used when no measured delay value is available for a crossing.
TEST_POINT_RADIUS_M	16, 18, 20 m	Radius used to generate candidate pedestrian test points around each intersection.
ARM_ANGLE_THRESHOLD_DEG	30°	Angular threshold used to group test points into intersection arms.
MAX_POINTS_PER_ARM	2	Maximum number of test points retained for each intersection arm.
MAX_ARMS	4	Maximum number of intersection arms considered in the route simulation.
ALERT_DELAY_THRESHOLD_S	30 s	Delay threshold used to count routes with alert conditions.
MIN_VALID_ROUTES	2	Minimum number of valid routes required to summarize an intersection.

The outputs generated in this stage are summarized in Table 4.5. The main output is the final intersection-level classification layer, while the remaining files are used mainly for debugging and validation.

Table 4.5: Outputs generated during the route simulation and classification stage.

Output	**Format**	**Content**
intersections_evaluated.geojson	GeoJSON	Final intersection-level classification layer containing delay indicators, route-based metrics, LOS class, severity score, alert status and visualization attributes.
intersections_test_points.geojson	GeoJSON	Debug layer containing the automatically generated test points used to simulate local pedestrian routes.
intersections_problematic.geojson	GeoJSON	Debug layer containing intersections where the simulation did not produce valid results or where processing issues were detected.
problematic_test_points.geojson	GeoJSON	Debug layer containing test points associated with problematic intersections.

Figure 4.9 shows the main spatial output generated in this stage. The figure is included here to illustrate the structure of the produced classification layer, while the interpretation of the spatial results is presented in Chapter 5.

Figure 4.9: Intersection-level classification layer generated at the end of the route simulation and classification stage. Source: Author

4.2.3 Zone Aggregation

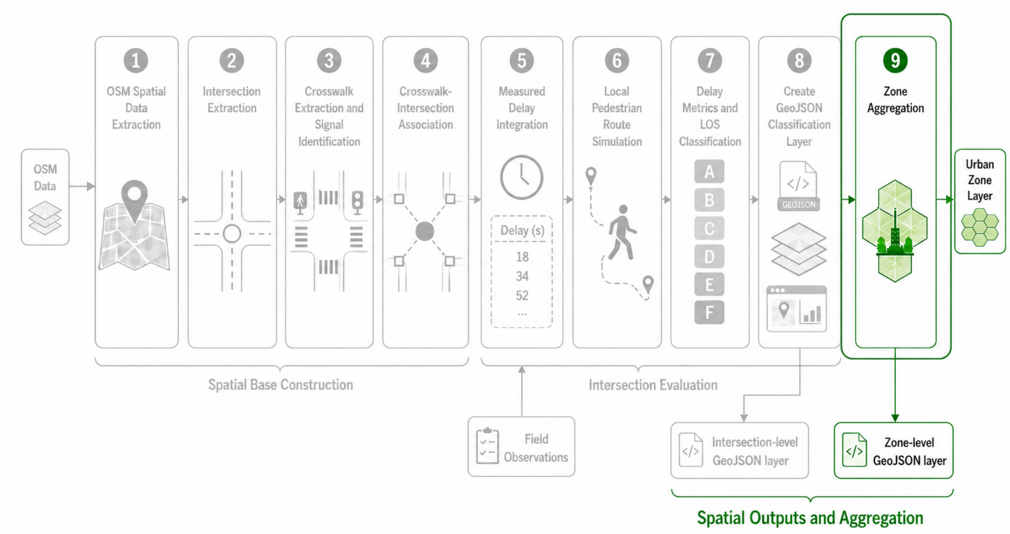

Figure 4.10: Workflow phase corresponding to spatial outputs and zone aggregation. Figure generated with the support of ChatGPT and manually reviewed by the author.

The third processing stage, highlighted in Figure 4.10, aggregates the intersection-level classification results by urban zones. While the previous stage produces one feature for each signalized intersection prepared for evaluation, this stage summarizes only the intersections with a valid LOS classification at a broader spatial scale. This makes it possible to represent pedestrian delay conditions not only at individual intersections, but also across larger urban areas.

The input data for this stage includes the `intersections_evaluated.geojson` layer produced in the previous stage and the urban zone layer used to divide the study area. Before the spatial association is performed, the zone layer is normalized by identifying the column that contains the zone name and by assigning a `zone_id`. The zone geometries are also validated to remove invalid or empty geometries before being used in spatial operations. These preparation steps are summarized in Algorithm 3, lines 1–4.

Both the intersection layer and the zone layer are projected to EPSG:3763 before the spatial join is carried out. This is done because the association between intersections and zones relies on metric distances. In the implementation, each intersection is associated with the nearest urban zone. The resulting layer stores, for each intersection, both its original classification attributes and the identifier and name of the corresponding zone. This spatial association process is represented in Algorithm 3, lines 5–7.

After the spatial join, the implementation prepares the LOS values for aggregation. The LOS field is first cleaned by removing extra spaces and converting the values to uppercase. Then, only intersections with a valid LOS class, from A to F, are retained for the zone-level calculations. This filtering step is important because the intersection-level layer also contains signalized intersections without measured delay values. These intersections are kept in the output layer for visualization and future data collection, but they do not represent valid delay-based evaluations and should not influence the zone scores. The cleaning, validation and score assignment steps are shown in Algorithm 3, lines 8–14.

For each valid intersection, the cleaned LOS class is converted into a numerical score using a mapping from A to F. In this mapping, LOS A corresponds to score 1 and LOS F corresponds to score 6. The zone-level indicators are then calculated by grouping the valid intersections by `zone_id` and `zone_name`. For each zone, the implementation calculates the average score, the worst score, the best score and the number of valid intersections used in the aggregation. The average score is used to assign the main zone-level LOS class, while the worst score is used to identify the worst LOS class present in the area. This aggregation logic is summarized in Algorithm 3, lines 15–20.

Zones without any valid LOS-classified intersection are kept in the output layer, but they are treated as zones without data. This means that their analytical classification fields remain empty or neutral, instead of being assigned an artificial good score. This distinction avoids interpreting default numerical delay values, such as zero, as valid low-delay results when no real delay measurement is available.

The calculated zone scores are then merged back into the original zone geometries. Each polygon keeps its geometry and receives the aggregated indicators as attributes. Visualization fields are also added, including colour, fill, stroke, stroke width, fill opacity and a specific field named `zone_color_rank`. This field is used to control the colour of the zone in the frontend. It assigns value 0 to zones without data and values 1 to 6 to zones classified from LOS A to LOS F. This separates the analytical metric, `avg_score`, from the categorical value used for map visualization. The merge and visualization steps are represented in Algorithm 3, lines 21–28.

The main logic of this stage is summarized in Algorithm 3.

Algorithm 3 Zone aggregation

Require: Evaluated intersection layer and urban zone layer
Ensure: Zone-level aggregation layer and zone score table

$I \leftarrow$ load `intersections_evaluated.geojson`
$Z \leftarrow$ load urban zone layer
normalize zone names and assign `zone_id`
validate zone geometries and export `zones_validated.geojson`
project I and Z to EPSG:3763
$J \leftarrow$ associate each intersection in I with the nearest zone in Z
reproject J to EPSG:4326
for all $i \in J$ **do**
    clean `los` value using string conversion, strip and uppercase
    $i.has\_valid\_los \leftarrow$ true if `los` is A, B, C, D, E, or F
    $i.score \leftarrow$ numerical LOS rank when `has_valid_los` is true
    set $i.score$ to null when `has_valid_los` is false
end for
$J_{zone} \leftarrow$ intersections in J with an associated `zone_name`
for all zone group z defined by `zone_id` and `zone_name` **do**
    calculate `avg_score`, `worst_score` and `best_score`
    count the number of associated intersections as `n_intersections`
    convert `avg_score` into `zone_los`
    convert `worst_score` into `worst_zone_los`
end for
merge the calculated indicators back into the original zone geometries
replace missing `n_intersections` values with 0
for all zone $z \in Z$ **do**
    clean and validate `zone_los`
    assign `zone_color_rank` from `zone_los`
    use rank 0 and no-data colour when `zone_los` is null
    assign `color`, `fill`, `stroke`, `stroke_width` and `fill_opacity`
end for
export `intersections_with_zone.geojson`
export `zones_with_scores.geojson`
export `zones_scores.csv`

The main parameters and configuration elements used in this stage are summarized in Table 4.6.

Table 4.6: Main parameters and configuration elements used in the zone aggregation stage.

Parameter	**Value**	**Purpose**
`METRIC_CRS`	EPSG:3763	Projected coordinate reference system used to perform the nearest-zone spatial association in metres.
`LOS_MAP`	A–F to 1–6	Conversion used to transform valid LOS classes into numerical scores for aggregation.
Valid LOS classes	A, B, C, D, E, F	Set of accepted LOS values used to decide which intersections are included in the zone-level aggregation.
`NO_DATA_COLOR`	Grey	Colour assigned to zones without any valid LOS-classified intersection.
`LOS_COLORS`	A–F colour scale	Colour mapping used to assign visualization colours to zones with valid aggregated classifications.
`zone_color_rank`	0–6	Numerical field used by the frontend to colour the zone layer, where 0 represents no data and 1–6 represent LOS A–F.
Zone name candidates	Multiple fields	List of possible attribute names used to identify the column containing the zone name.

The outputs generated in this stage are summarized in Table 4.7.

Table 4.7: Outputs generated during the zone aggregation stage.

Output	**Format**	**Content**
`zones_validated.geojson`	GeoJSON	Validated urban zone layer with normalized zone identifiers and zone names.
`intersections_with_zone.geojson`	GeoJSON	Intersection-level layer with the associated zone identifier and zone name.
`zones_with_scores.geojson`	GeoJSON	Final zone-level aggregation layer containing average, worst and best scores, number of valid intersections, zone-level LOS, worst zone-level LOS, `zone_color_rank` and visualization attributes.
`zones_scores.csv`	CSV	Tabular summary of the calculated zone-level indicators.

Figure 4.11 shows the zone-level spatial output generated in this stage. The figure is included here only to illustrate the aggregated layer produced by the implementation; its results are analysed in Chapter 5.

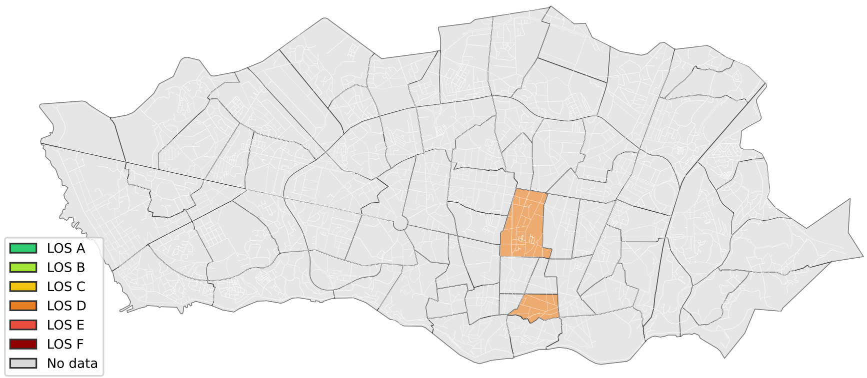

Figure 4.11: Zone-level aggregation layer generated from the valid intersection-level classification results. Source: Author

4.3 Conclusion

This chapter described the technical implementation of the proposed methodology. The workflow was implemented in Python using geospatial, tabular and network analysis libraries, with EPSG:3763 used for metric operations and EPSG:4326 used for exporting the final spatial layers.

The implementation was organized into three main stages. The first stage extracted and prepared the spatial base from OSM, including the road network, consolidated intersections, pedestrian crossings and signalized crossing information. The second stage used these outputs to build local pedestrian graphs, generate test points, simulate pedestrian routes, calculate delay indicators and create the final intersection-level classification layer. The third stage aggregated the classified intersections by urban zones, producing a zone-level spatial layer.

Throughout the implementation, intermediate outputs were generated to support validation and inspection of the process. These outputs made it possible to check the extracted intersections, crossing associations, simulated test points, evaluated intersections and aggregated zones separately. The final GeoJSON layers produced by the workflow provide the basis for the analysis presented in the following chapter.

Chapter 5

Results and Discussion

This chapter presents the final spatial outputs produced by the implemented methodology. While Chapter 4 described how the workflow was built, this chapter focuses on how its results are displayed and interpreted in the urban monitoring platform used by the city of Porto.

The first part presents the point-based result, including its visualization logic and the attributes available for each evaluated location. The second part moves to the area-level output, where valid classifications are aggregated by urban area. The chapter ends with a discussion of how these two views should be read together.

5.1 Intersection-Level Layer

The point-based output is the main result of the workflow at the local scale. It brings together the methodological steps described in Chapter 3 and the implementation process described in Chapter 4, translating them into a spatial product that can be visualized in the urban monitoring platform used by the city of Porto. In the platform, this result appears as *Classificação de Interseções*, which can be translated as *Intersection Classification*.

The current output contains 299 locations prepared for evaluation. These do not represent every junction in the city, but the subset retained by the workflow because of their association with signalized pedestrian crossings, as described in Chapter 3. This is the group where the proposed method can be applied and where delay-based classification can be produced when measured values are available.

The map also separates places with manually measured delay data from those where field information is still missing. Coloured markers correspond to locations where these measured values were available and where the LOS classification could be calculated. Grey markers represent signalized crossing areas that were identified and prepared by the workflow, but that do not yet have measured delay values. These grey markers should not be interpreted as good or bad results. They simply show where the assessment is still incomplete.

Figure 5.1 shows this output in its final form. The map gives a city-wide view of the signalized locations covered by the workflow, while making the already evaluated subset visible through the coloured markers.

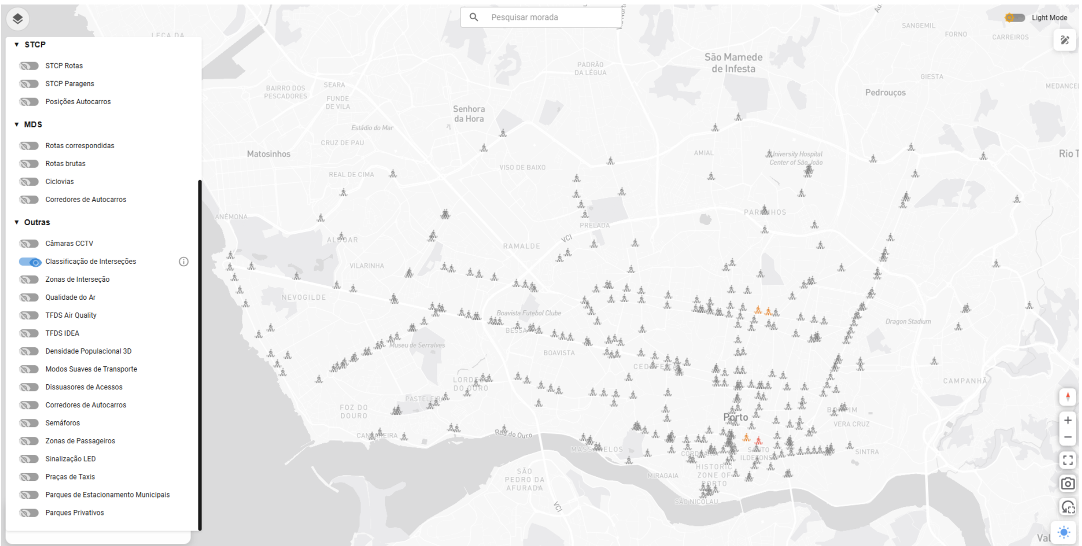

Figure 5.1: Point-based classification output displayed in the urban monitoring platform. Coloured markers represent locations with measured delay values, while grey markers represent locations without field delay data. Source: Author

Each point can also be inspected individually. Figure 5.2 shows an example of the information displayed for one evaluated case. Since the platform interface is in Portuguese, the main fields are described here in English. *Classificação de Interseção* means *Intersection Classification*. The field *ID* identifies the evaluated location. *LOS* shows the assigned Level of Service. *Atraso Médio* and *Atraso Máximo* correspond to the average and maximum pedestrian delay. *Rotas Testadas* means the number of simulated routes and *Rotas em Alerta* indicates how many of those routes were classified as alert routes. Finally, *Severidade* corresponds to the assigned severity score.

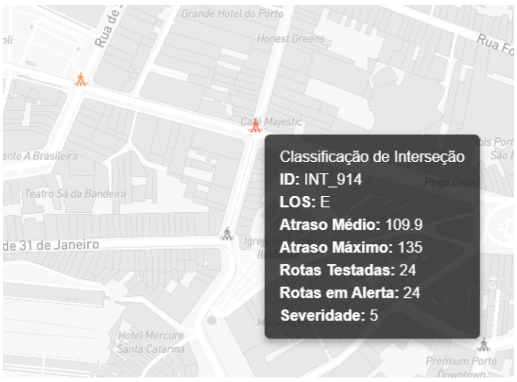

Figure 5.2: Example of the attribute information associated with one evaluated location in the urban monitoring platform. Source: Author

The example shown in Figure 5.2 corresponds to INT_914. This location was classified as

LOS E, with an average delay of 109.9 s and a maximum delay of 135 s. The platform also shows that 24 pedestrian routes were tested and that all of them were classified as alert routes. The severity score is 5. These values indicate high pedestrian delay across the simulated movements considered in the evaluation.

This attribute view connects directly with the route simulation stage described in Chapter 4. The user does not only see a coloured marker; the popup also shows the delay values, the number of analysed routes and the alert information behind the classification. This makes the result easier to interpret for both technical users and urban management teams.

Overall, this spatial output turns the methodology into a practical result. It shows the locations already evaluated from field data, while keeping the remaining signalized crossing areas visible for future data collection. As more delay measurements are added, the same structure can be updated and used to expand the pedestrian mobility assessment across more places.

5.2 Zone-Level Visualization

The area-level output complements the point-based result presented in the previous section. While the previous map focuses on individual locations, this view summarizes the same assessment across broader urban areas. As described in Chapter 4, this output is created by associating valid classifications with the corresponding spatial units of Porto and calculating aggregated indicators.

In the urban monitoring platform used by the city of Porto, this result appears as *Zonas de Interseção*, which can be translated as *Intersection Zones*. Each polygon represents one urban area and its colour reflects the aggregated LOS obtained from the valid locations associated with it. This view can also be displayed together with the point markers, making it easier to relate the broader area-based result with the local information behind it.

Figure 5.3 shows this combined view. The coloured polygons represent areas where the available manually measured delay data was enough to produce an aggregated result. Grey polygons correspond to areas where this was not yet possible. When both outputs are displayed together, the platform supports two levels of reading: a general spatial summary and a more detailed inspection of specific locations.

Figure 5.3: Area-level output displayed together with the point-based classification result in the urban monitoring platform. Source: Author

Each area can also be inspected through its attributes. Since the platform interface is in Portuguese, the main fields are described here in English. *Informação de Zona* means *Zone Information*. *Nome da Zona* identifies the name of the area. *LOS da Zona* is the aggregated LOS. *Score Médio* corresponds to the average score, while *Pior Score* and *Melhor Score* correspond to the worst and best scores among the associated classified locations. *Nº de Interseções* indicates how many locations are linked to that area and *Pior LOS* shows the worst LOS found within that set.

Figures 5.4 and 5.5 show two classified examples. Both *Batalha* and *Camões* are classified as LOS D and therefore appear with the same polygon colour. This is expected, since the colour represents the final area-level class. However, the attribute values show that the two cases are not identical internally.

In *Batalha*, the area has an average score of 4.5, a worst score of 5, a best score of 4 and eight associated classified locations. The worst LOS inside this area is E. In *Camões*, the average score is 4.0, with a worst score of 4, a best score of 4 and fifteen associated classified locations. In this case, the worst LOS is D. This comparison shows why it is useful to include both average and worst values in the popup. The map colour gives a quick reading of the final class, while the attributes provide more detail about the internal conditions. In this case, *Batalha* is slightly more critical than *Camões*, even though both share the same final LOS.

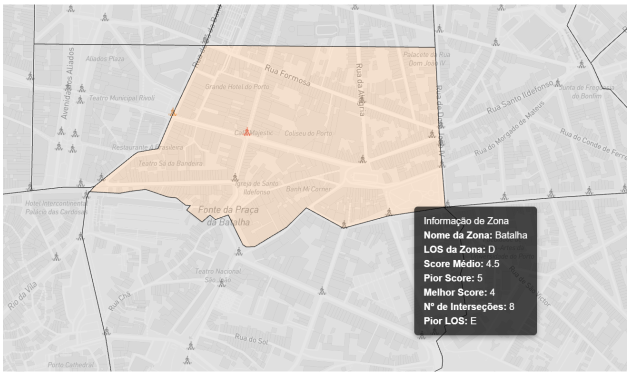

Figure 5.4: Attribute information for the area *Batalha* in the urban monitoring platform. Source: Author

Figure 5.5: Attribute information for the area *Camões* in the urban monitoring platform. Source: Author

It is also important to show how areas without enough valid information appear in the platform. Figure 5.6 presents the example of *Pinheiro Manso*. In this case, the polygon remains grey and the main analytical fields appear as *N/A*. This means that, although the area exists in the spatial output and can still be inspected, the available information was not enough to produce a valid aggregated classification. Showing these cases explicitly makes the current data coverage clear.

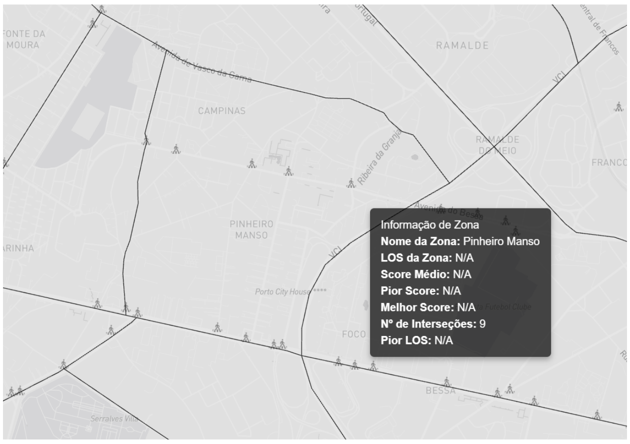

Figure 5.6: Example of an area without a valid aggregated classification result in the urban monitoring platform. Source: Author

Taken together, these examples show that the area-level result should be read as a summary of the point-based classifications rather than as a replacement for them. The colour of each polygon provides a quick overview of the aggregated LOS, but the attributes are needed to understand the values behind that class. The local map remains necessary to inspect the specific conditions behind each aggregated result.

5.3 Pedestrian Mobility within Smart City Systems

The results presented in this chapter can also be read within a broader smart city context. Although the implemented workflow is based on open geospatial data and manually measured pedestrian delay, the final outputs are structured in a way that can be connected to urban monitoring systems, sensing infrastructure and future data sources. This is important because pedestrian mobility is part of smart mobility and Intelligent Transportation Systems (ITS), even if it is often less visible than vehicle traffic in operational platforms.

The work developed in this dissertation contributes to this context by transforming pedestrian delay into structured spatial information. The point-based result makes it possible to inspect specific signalized locations, while the area-level output summarizes valid classifications across broader urban areas. These two views allow pedestrian conditions to be represented in the same type of platform environment where other urban indicators are already displayed. In that sense, the results can be read as part of a data pipeline that collects, processes, classifies and visualizes information for urban decision support.

This connection is also relevant from an electrical engineering perspective. A future version of the system could be supported by different types of sensing technologies. Cameras could be used to estimate pedestrian counts, waiting times or crossing behaviour through computer vision methods. Wi-Fi, Bluetooth or BLE sensing could help estimate presence, movement patterns or dwell time from wireless signals emitted by devices. Pedestrian push buttons, infrared sensors, radar sensors or other embedded devices could also provide local information about crossing demand.

Wearable devices and smartphone-based data could contribute with information about walking routes, travel times or exposure to environmental conditions, although these sources would require careful attention to privacy and representativeness. These components are consistent with Internet of Things (IoT)-based smart city architectures, where sensing, communication and data processing are combined to support urban services [48].

These technologies would not replace the geospatial workflow developed in this dissertation. Instead, they could complement it. The current approach already defines where the information should be stored, how delay values can be associated with signalized crossings, how local route simulation can use those values and how the final classification can be displayed. Sensor-based data could therefore feed the same structure with more frequent or more detailed measurements. For example, automatically estimated waiting times could update the delay fields, pedestrian counts could help prioritize locations with higher demand and signal timing data could improve the interpretation of the LOS results.

The use of embedded systems and edge computing could be particularly useful in this type of application. Instead of sending all raw data to a central platform, part of the processing could be done near the sensor. A camera installed close to a crossing could, for example, count pedestrians or estimate waiting time locally, sending only aggregated values to the urban monitoring platform. This would reduce bandwidth requirements and avoid the unnecessary transmission of sensitive raw data. Edge and cloud computing are commonly discussed in smart city systems because they allow processing tasks to be distributed between local devices, intermediate nodes and central platforms [49]. Edge Artificial Intelligence (AI) could also support faster processing and make the system more scalable if several locations were monitored at the same time [50].

Communication networks would also play an important role in a more dynamic implementation. Different sensors may require different communication technologies depending on the amount of data, latency, energy consumption and coverage needed. A camera-based system may require higher bandwidth, while a simple presence detector could use a lower-power communication solution. These technical choices are relevant because pedestrian monitoring in public space needs to be reliable, maintainable and feasible at city scale.

Energy consumption is another significant aspect. Sensors installed in urban space may need to operate for long periods and under outdoor conditions. Some devices may be connected to the electrical grid, while others may depend on batteries or small energy harvesting systems. Low-power embedded processing and efficient communication protocols can therefore become important design requirements. These constraints show that a future sensor-based version of the work would not be only a data problem, but also a hardware, communication and energy management problem.

At the same time, any future integration of sensing technologies would need to consider privacy and data protection from the beginning. Pedestrian mobility data can reveal movement patterns, routines and the presence of people in specific places. Camera-based systems may capture identifiable images, while wireless sensing and wearable data may involve device identifiers or location traces. For this reason, privacy-preserving approaches would be essential. These could include local processing, anonymization, aggregation before transmission, limited data retention and the avoidance of raw image storage whenever possible. In the European context, this would need to be aligned with GDPR principles.

This discussion shows how the dissertation connects with smart city systems and with several areas of electrical and computer engineering. The current implementation is static and based on available data, but it prepares a structure that could later receive information from IoT devices, embedded sensing units, communication networks and edge processing systems. The geospatial outputs presented in this chapter can therefore be seen as the visible part of a wider possible architecture for pedestrian mobility monitoring.

In this sense, the work contributes to making pedestrian mobility more present in urban data systems. By representing delay conditions in a platform-ready format, the dissertation helps bring the pedestrian experience into the same digital environment where other city systems are monitored.

5.4 Discussion

The results presented in this chapter show how the proposed workflow can translate pedestrian delay into information that is both spatial and interpretable. The point-based view provides the most detailed reading, since each marker stores delay indicators, route-based metrics, LOS class, severity score and alert information for a specific signalized location. This makes it possible to inspect individual cases and understand the values behind each classification.

The distinction between coloured and grey markers is important for interpreting the map correctly. Coloured markers represent places with manually measured delay values and a valid LOS classification. Grey markers represent signalized crossing areas that were identified and prepared by the workflow, but where field delay data is not yet available. This means that the map shows both the current evaluated subset and the locations where the assessment structure is already prepared, but still incomplete.

The aggregated view provides a broader reading of the same information. By grouping valid classifications by urban area, it becomes easier to compare different parts of the city and identify where pedestrian delay conditions may require more attention. However, this view should be understood as a summary. It does not replace the local map, because average values can hide differences within the same area.

The examples of *Batalha* and *Camões* illustrate this point. Both areas have the same final classification, LOS D and therefore share the same polygon colour. Still, their internal values are different. *Batalha* has a higher average score and includes at least one location classified as LOS E, while *Camões* has a lower average score and no associated location worse than LOS D. This shows why the popup includes both average and worst values. The average score gives the general classification of the area, while the worst score helps identify whether a more critical local condition exists inside it.

The example of *Pinheiro Manso* shows the opposite case. Since this area does not contain valid LOS-classified locations, it is represented in grey and its analytical fields appear as *N/A*. This distinction avoids interpreting missing data as a positive result. A grey polygon does not mean that pedestrian conditions are good or bad; it means that there is not enough valid information to classify it.

The results also show why this type of assessment fits within a smart city context. The current implementation is based on open geospatial data and manually measured delay values, but the same structure could later receive information from sensors, embedded devices or other urban data sources. In that sense, the platform-ready outputs produced in this dissertation can be understood as the visible part of a broader data pipeline, where pedestrian information is collected, processed, classified and displayed for urban decision support.

Overall, the two spatial views work best when used together. The aggregated map supports a general reading of pedestrian delay conditions, while the local result keeps the detail needed to interpret each classification. This combination is more useful than a single table of values or an isolated set of field observations, because the results can be inspected directly in their spatial context and connected to future smart mobility applications.

Chapter 6

Conclusion

This dissertation addressed the assessment of pedestrian delay conditions at signalized urban crossings in Porto. The work was motivated by the fact that pedestrian mobility is often less visible in urban monitoring tools than other transport or infrastructure systems, even though walking is part of most daily trips and is strongly affected by local conditions at crossings.

The main goal was to understand how geospatial data and measured pedestrian delay could be combined to assess pedestrian conditions at signalized locations and represent the results in an urban monitoring platform. To answer this question, the dissertation developed a data-driven workflow that connects open spatial data, field observations, local route simulation, delay-based LOS classification and platform-ready geospatial outputs.

The workflow starts from OSM data, which is used to extract the road network, identify relevant urban junctions, detect pedestrian crossings and distinguish those with pedestrian signal control. Field observations are then linked to selected crossings, allowing measured waiting times to be introduced into the spatial structure. These values are used during the route simulation stage, where possible pedestrian movements are tested around each evaluated location.

The simulated routes make it possible to calculate delay indicators and assign a LOS class to each evaluated case. The results are then exported in GeoJSON format, with the attributes needed for visualization and inspection in the urban monitoring platform used by the city of Porto. A second spatial output aggregates the valid classifications by urban area, providing a broader view of the same assessment.

6.1 Main Outcomes

The work produced two main spatial outputs. The first is a point-based result, where each evaluated location stores delay values, route-based metrics, LOS class, severity score, alert information and visualization attributes. This output allows individual signalized locations to be inspected directly on a map and makes the values behind each classification visible.

The second output summarizes the valid classifications by urban area. This broader view makes it easier to compare different parts of the city and identify where pedestrian delay conditions may require more attention. Areas without valid classified locations are kept in the output but represented as no data, avoiding the mistake of interpreting missing measurements as favourable pedestrian conditions.

Together, these two outputs show the importance of combining detailed and aggregated views. The point-based result keeps the local information needed to understand specific cases, while the area-level view helps communicate broader spatial patterns. This combination is useful because pedestrian delay is local by nature, but urban management often also needs summaries that can be read at a wider scale.

Another important outcome is the preparation of the results for integration into an urban monitoring context. The final files are not only analytical outputs; they are structured with identifiers, classification fields, scores, colours and visualization attributes. This makes them easier to reuse, update and display in a platform environment.

As part of the dissemination of this work, a conference paper entitled "Assessing Pedestrian Mobility Patterns in Crosswalks through Data-driven Geospatial Modelling" was submitted to the 2026 IEEE International Smart Cities Conference (ISC2). The paper presents the proposed data-driven geospatial framework for pedestrian-intersection assessment and discusses its application to the city of Porto. At the time of writing, the submission is still under review, so no publication outcome is yet available.

Overall, the dissertation shows that pedestrian delay can be transformed from isolated field measurements into structured spatial information. Even with limited measured data, the proposed workflow creates a clear base that can be expanded as more information becomes available.

6.2 Limitations and Future Work

The main limitation of this work is the small number of manually measured delay values. Since field observations were collected only for a limited set of signalized crossings, the current result should not be interpreted as a complete diagnosis of pedestrian delay conditions in Porto. Instead, it should be understood as a first operational version of the assessment structure. Future work should increase the number of measured locations, allowing the classification to cover a larger part of the city.

A second limitation is related to the quality and completeness of OSM data. Although this source provides a useful spatial base, pedestrian crossings, traffic signals and network elements may be mapped with different levels of detail. Some features may be missing, represented with inconsistent tags, or stored using different geometry types. Future versions of the workflow could benefit from validation with municipal data or additional field checks.

The route simulation also simplifies real pedestrian behaviour. The generated test points and local pedestrian graph allow possible movements to be tested, but they do not represent actual trajectories or individual decisions. The method does not capture pedestrian volumes, crowding, perceived safety, comfort, accessibility constraints, or weather exposure. Future work could improve this stage by using observed pedestrian flows, origin-destination data, or more detailed pedestrian network information.

Another limitation is that the LOS classification is based only on delay. Waiting time is a relevant and measurable indicator for signalized crossings, but it does not describe the full walking experience. A location with low delay may still be uncomfortable or unsafe due to traffic speed, poor visibility, narrow sidewalks, or other local conditions. Future versions could include additional indicators, such as crossing width, number of crossing stages, pedestrian volume, sidewalk quality, accessibility barriers, or perceived safety.

The workflow could also be extended with more dynamic data sources. Pedestrian sensors, cameras, Wi-Fi or BLE detections, pedestrian counts and traffic signal systems could provide more frequent information about pedestrian presence, waiting times and crossing behaviour. If these sources become available, the same structure could support regular updates and a more dynamic form of pedestrian mobility monitoring.

Finally, the integration with the urban monitoring platform could be developed further. Future versions could include automatic updates, temporal comparisons, dashboards, or alert mechanisms. This would allow the assessment to move from a static spatial output to a more operational tool for supporting pedestrian-oriented decisions.

6.3 Final Remarks

This dissertation provides a first structured approach for representing pedestrian delay at signalized urban crossings in Porto. The current result is limited by the available field data, but the workflow creates a base that can grow as more information is collected.

The main value of the work lies in connecting different parts of the assessment process into a single spatial structure. Open geospatial data, measured delay values, route simulation, classification and platform-ready outputs are brought together in a way that can be inspected, updated and expanded.

By doing this, the dissertation contributes to making pedestrian mobility more visible in urban data systems. It does not aim to solve every aspect of walkability, but it shows how one specific part of the pedestrian experience, waiting time at signalized crossings, can be measured, classified and represented in a form that supports future urban monitoring.

References

[1] D. G. Costa, J. C. N. Bittencourt, F. Oliveira, J. P. J. Peixoto, and T. C. Jesus, "Achieving sustainable smart cities through geospatial data-driven approaches," *Sustainability*, vol. 16, no. 2, p. 640, 2024. DOI: 10.3390/su16020640.

[2] S. Goumiri, S. Yahiaoui, and S. Djahel, "Smart mobility in smart cities: Emerging challenges, recent advances and future directions," *Journal of Intelligent Transportation Systems*, vol. 29, no. 1, pp. 81–117, 2025.

[3] S. Barns, "Smart cities and urban data platforms: Designing interfaces for smart governance," *City, Culture and Society*, vol. 12, pp. 5–12, 2018. DOI: 10.1016/j.ccs.2017.09.006.

[4] OECD, *Smart city data governance: Challenges and the way forward*, 2023. DOI: 10.1787/e57ce301-en.

[5] K. Dovey and E. Pafka, "What is walkability? the urban dma," *Urban Studies*, vol. 57, pp. 93–108, Jan. 2020. DOI: 10.1177/0042098018819727.

[6] H. Chaudhary, S. Mahmood, Y. Su, M. S. Mehmood, and M. I. Ahamad, "Urban walkability assessment: Methods and implications for city planning," *Environment, Development and Sustainability*, pp. 1–31, 2025.

[7] F. Moura, P. Cambra, and A. Gonçalves, "Measuring walkability for distinct pedestrian groups with a participatory assessment method: A case study in lisbon," *Landscape and Urban Planning*, vol. 157, pp. 282–296, Jan. 2017. DOI: 10.1016/j.landurbplan.2016.07.002.

[8] Transportation Research Board, *Highway Capacity Manual*. National Academies Press, 2016.

[9] S. Basu et al., "Understanding pedestrian waiting experiences at signalized intersections," in *Proceedings of the ACM on Human-Computer Interaction*, ACM, 2024. DOI: 10.1145/3748636.3762798.

[10] N. M. Zafri, A. I. Rony, and N. Adri, "Analysis of pedestrian crossing speed and waiting time at intersections in dhaka," *Infrastructures*, vol. 4, no. 3, p. 39, 2019. DOI: 10.3390/infrastructures4030039.

[11] S. Marisamynathan and P. Vedagiri, "Pedestrian perception-based level-of-service model at signalized intersection crosswalks," *Journal of Modern Transportation*, vol. 27, pp. 266–281, 2019. DOI: 10.1007/s40534-019-00196-5.

[12] Porto Digital. "Foundations of porto's urban platform." Accessed: 2026-05-29. [Online]. Available: https://portodigital.pt/w/plataforma-urbana-do-porto.

[13] Associação Porto Digital, *Porto Digital*, https://portodigital.pt/, Accessed: 2026-06-05, 2026.

[14] J. F. Sallis et al., "Use of science to guide city planning policy and practice: How to achieve healthy and sustainable future cities," *The Lancet*, vol. 388, no. 10062, pp. 2936–2947, 2016. DOI: 10.1016/S0140-6736(16)30068-X.

[15] A. Ferdman, "Walking and its contribution to objective well-being," *Journal of Planning Education and Research*, Sep. 2019. DOI: 10.1177/0739456X19875195.

[16] M.-B. Fernandez Nunez, M. Maciejewska, L. Mojica, and O. Marquet, "Car-use reduction in 15-minute cities. a matter of modal shift or shorter travel distances?" *Journal of Urban Mobility*, vol. 6, p. 100 093, 2024. DOI: 10.1016/j.urbmob.2024.100093.

[17] H. Butler, M. Daly, A. Doyle, S. Gillies, T. Schaub, and S. Hagen, *The geojson format*, RFC 7946, 2016. DOI: 10.17487/RFC7946.

[18] G. Boeing, "Osmnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks," *Computers, Environment and Urban Systems*, vol. 65, pp. 126–139, 2017. DOI: 10.1016/j.compenvurbsys.2017.05.004.

[19] G. Boeing, "Modeling and analyzing urban networks and amenities with osmnx," *Environment and Planning B: Urban Analytics and City Science*, 2024.

[20] A. Bartzokas-Tsiompras, "Utilizing openstreetmap data to measure and compare pedestrian street lengths in 992 cities around the world," *European Journal of Geography*, vol. 13, no. 2, pp. 127–141, 2022.

[21] A. Qu, A. Valiveru, C. Tang, V. Jayawardana, B. Freydt, and C. Wu, "What is a typical signalized intersection in a city? a pipeline for intersection data imputation from openstreetmap," *arXiv preprint arXiv:2405.13480*, 2024. DOI: 10.48550/arXiv.2405.13480.

[22] B. G. Pinto and P. Chamusca, "The 15-minute city in portugal: Reality, aspiration, or utopia?" *Urban Science*, vol. 9, no. 9, 2025, ISSN: 2413-8851. DOI: 10.3390/urbansci9090330. [Online]. Available: https://www.mdpi.com/2413-8851/9/9/330.

[23] N. Abuwaer, S. Ullah, and S. G. Al-Ghamdi, "Establishing the nexus between urban walkability and thermal comfort in a changing climate," *Nature Cities*, vol. 2, no. 9, pp. 801–811, 2025.

[24] R. Prieto-Curiel and F. Barroso, "The growing dominance of cars in suburban areas," *Transportation Research Interdisciplinary Perspectives*, vol. 32, p. 101 559, 2025, ISSN: 2590-1982. DOI: https://doi.org/10.1016/j.trip.2025.101559.

[25] F. Fonseca et al., "Perceived walkability and respective urban determinants: Insights from bologna and porto," *Sustainability*, vol. 14, no. 15, p. 9089, 2022. DOI: 10.3390/su14159089.

[26] X. Guo et al., "Evaluating the quality of high-frequency pedestrian commuting streets: A data-driven diagnostic framework," *Smart Cities*, vol. 8, no. 3, 2025.

[27] K. Shoji, B. Tarver, and A. Sonta, "Data-driven insights on urban walkability: Modeling the relationships between urban form and pedestrian activity," *Cities*, vol. 175, p. 107 148, 2026, ISSN: 0264-2751. DOI: https://doi.org/10.1016/j.cities.2026.107148.

[28] Z. Yang, Z. Gong, Q. Zhang, and J. Wang, "Analysis of pedestrian-related crossing behavior at intersections: A latent dirichlet allocation approach," *International Journal of Transportation Science and Technology*, vol. 12, no. 4, pp. 1052–1063, 2023. DOI: 10.1016/j.ijtst.2022.12.003.

[29] K. Papandreou et al., “Improving pedestrian level of service at signalized intersections through dynamic signal timing,” *Transportation Research Part C*, pp. 125–140, 2017.

[30] T. Ahmed et al., “A new pedestrian crossing level of service method for promoting safe pedestrian crossing in urban areas,” *International Journal of Environmental Research and Public Health*, vol. 18, no. 16, p. 8813, 2021. DOI: 10.3390/ijerph18168813.

[31] H. Zhang, D. Zhang, and R. Wang, “Evaluation of pedestrian level of service at signalised intersections from the elderly perspective,” *Promet - Traffic&Transportation*, vol. 35, no. 3, pp. 434–445, 2023. DOI: 10.7307/ptt.v35i3.130.

[32] R. Nowar and M. I. Burke, “What are the most important factors for pedestrian level-of-service estimation? a systematic review of the literature,” *Transportation Research Record*, vol. 2672, no. 35, pp. 101–117, 2018. DOI: 10.1177/0361198118790623.

[33] M. Jabbari et al., “The pedestrian network concept: A systematic literature review,” *Journal of Urban Mobility*, vol. 3, p. 100 051, 2023, ISSN: 2667-0917. DOI: https://doi.org/10.1016/j.urbmob.2023.100051.

[34] A. Dhanani, L. Tarkhanyan, and L. Vaughan, “Estimating pedestrian demand for active transport evaluation and planning,” *Transportation Research Part A: Policy and Practice*, vol. 103, pp. 54–69, 2017. DOI: 10.1016/j.tra.2017.05.020.

[35] D. Tan, M. Younis, W. Lalouani, S. Fan, and G. Song, “A novel pedestrian road crossing simulator for dynamic traffic light scheduling systems,” *Journal of Intelligent Transportation Systems*, 2024. DOI: 10.1080/15472450.2023.2186229.

[36] A. Kurkcu and K. Ozbay, “Estimating pedestrian densities, wait times, and flows with wi-fi and bluetooth sensors,” *Transportation Research Record*, vol. 2644, no. 1, pp. 72–82, 2017. DOI: 10.3141/2644-09.

[37] V.-D. Stanciu, “Privacy-friendly wifi-based crowd monitoring for pedestrian dynamics analytics,” Ph.D. dissertation, University of Twente, 2022. DOI: 10.3990/1.9789036554916.

[38] Câmara Municipal de Lisboa, *Lisboa Inteligente*, https://lisboaaberta.cm-lisboa.pt/index.php/pt/lisboa-inteligente, Accessed: 2026-06-29, 2026.

[39] L. Chen et al., “Estimating vehicle and pedestrian activity from town and city traffic cameras,” *Sensors*, vol. 21, no. 14, p. 4565, 2021. DOI: 10.3390/s21144565.

[40] L. E. Navarro-Serment, N. Lerner, D. Ratté, and G. Holness, “Monitoring and predicting pedestrian behavior using traffic cameras,” U.S. Department of Transportation, Tech. Rep., 2018.

[41] H. Liu, V. V. Gayah, and M. W. Levin, “A max pressure algorithm for traffic signals considering pedestrian queues,” *Transportation Research Part C: Emerging Technologies*, vol. 169, p. 104 865, 2024. DOI: 10.1016/j.trc.2024.104865.

[42] TM4CAD Consortium, “Implementation aspects of distributed odd attribute value awareness,” CEDR Transnational Road Research Programme, Tech. Rep., 2021, Accessed: 2026-06-12. [Online]. Available: https://tm4cad.project.cedr.eu/deliverables/TM4CAD_D4.1_final.pdf.

[43] Z. Luo, A. M. Molan, J. E. Hummer, A. Pande, and S. Osafo-Gyamfi, “Pedestrian safety and efficiency at emerging alternative intersection designs with three-phase traffic signals,” *Innovative Infrastructure Solutions*, vol. 10, no. 1, 2025. DOI: 10.1007/s41062-024-01841-7.

[44] European Parliament and Council of the European Union, *Regulation (eu) 2016/679 of the european parliament and of the council: General data protection regulation*, Official Journal of the European Union, 2016.

[45] Gabinete de Estratégia e Estudos, "NUTS III: Área Metropolitana do Porto," Ministério da Economia e Coesão Territorial, Tech. Rep., 2023, Source: INE (AER), except when specified otherwise.

[46] CESOP-Local. "Porto municipality," Accessed: May 29, 2026. [Online]. Available: https://cesop-local.ucp.pt/en/autarchy/porto-municipality.

[47] Topographic Map. "Porto topographic map, elevation, terrain." Accessed: 2026-05-29. [Online]. Available: https://en-ie.topographic-map.com/map-78dxt6/Porto/.

[48] A. S. Syed, D. Sierra-Sosa, A. Kumar, and A. Elmaghraby, "Iot in smart cities: A survey of technologies, practices and challenges," *Smart Cities*, vol. 4, no. 2, pp. 429–475, 2021. DOI: 10.3390/smartcities4020024.

[49] M. Trigka and E. Dritsas, "Edge and cloud computing in smart cities," *Future Internet*, vol. 17, no. 3, p. 118, 2025. DOI: 10.3390/fi17030118.

[50] K. S. Velaga et al., "Edge ai for smart cities: Foundations, challenges, and opportunities," *Smart Cities*, vol. 8, no. 6, p. 211, 2025. DOI: 10.3390/smartcities8060211.