UNIVERSITY
INSTITUTE
OF LISBON

Urban Tourism Crowding Dynamics: Carrying Capacity and Digital Twinning

Duarte Sampaio Belchior de Almeida

Master in **Computer Science and Business Management**

Supervisor:
PhD Fernando Brito e Abreu, Associate Professor,
Iscte

November, 2021

[This page has been intentionally left blank]

TECHNOLOGY
AND ARCHITECTURE

Urban Tourism Crowding Dynamics: Carrying Capacity and Digital Twinning

Duarte Sampaio Belchior de Almeida

Master in **Computer Science and Business Management**

Supervisor:
PhD Fernando Brito e Abreu, Associate Professor,
Iscte

November, 2021

[This page has been intentionally left blank]

Urban Tourism Crowding Dynamics:
Carrying Capacity and Digital Twinning

[This page has been intentionally left blank]

To my family and friends, for their unconditional support and constant encouragement.

[This page has been intentionally left blank]

Acknowledgements

I want to express my sincere gratitude to my dissertation supervisor Prof. Dr. Fernando Brito e Abreu, for believing in my capabilities, and for his guidance, advising and constant support. I would also like to thank my family, friends and colleagues for all their support and motivation. Special thanks to Doctor Patrick Taillandier, one of the main GAMA platform developers, for his invaluable help in understanding the platform and solving related issues.

This work was produced with the support of INCD funded by FCT and FEDER under the project 01/SAICT/2016 nº 022153. I am particularly grateful to INCD's technical team for their timely support in providing a cloud-based solution (computing and storage) tailored to the needs of my research project.

Lisbon, November, 2021

Duarte Sampaio Belchior de Almeida

[This page has been intentionally left blank]

Resumo

O aumento da atividade turística a nível global tem levado à superlotação, causando danos aos ecossistemas locais e degradação da experiência turística. Para planear a atividade turística é necessário definir indicadores adequados e entender as dinâmicas das multidões turísticas.

Os principais objetivos desta dissertação são o desenvolvimento de (1) um algoritmo para avaliar a capacidade de carga física (CCF) de fino grão espacial para uma malha urbana complexa, (2) um modelo de simulação baseado em agentes para o escoamento de participantes em eventos de atração turística em espaços abertos e (3) um modelo de simulação baseado em agentes usando o algoritmo de CCF para análise do stress de aglomeração de turistas em cenários de malha urbana restritiva.

Os dados abertos do OpenStreetMap foram usados nesta investigação. O algoritmo CCF proposto foi testado na freguesia de Santa Maria Maior, em Lisboa, que tem uma malha urbana antiga e complexo. A plataforma GAMA baseada em agentes foi usada nos dois estudos de simulação. O primeiro comparou dois cenários (normal e COVID-19) em três grandes espaços públicos de Lisboa e o segundo analisou o stress de aglomeração causado pela chegada de navios ao Terminal de Cruzeiros de Lisboa.

Os resultados mostram a viabilidade do algoritmo proposto para determinar a CCF de zonas com tecidos urbanos complexos e a sua aplicação como valor de referência inicial para a avaliação do stress de superlotação em tempo real, nomeadamente na avaliação de cenários de aglomeração turística excessiva, tanto em espaços abertos, como em malhas urbanas intrincadas.

Palavras-chave: capacidade de carga; modelação baseada em agentes; turismo; superlotação; aglomeração turística

[This page has been intentionally left blank]

Abstract

The increase in tourism activity globally has led to overcrowding, causing damage to local ecosystems and degradation of the tourism experience. To plan tourist activity it is necessary to define adequate indicators and understand the dynamics of tourist crowds.

The main goals of this dissertation are the development of (1) an algorithm for assessing spatially fine-grained, physical carrying capacity (PCC) for a complex urban fabric, (2) an agent-based simulation model for the egress of participants in public open space tourism attraction events and (3) an agent-based simulation model using the PCC algorithm for tourism crowding stress analysis in urban fabric constrained scenarios.

OpenStreetMap open-data was used throughout this research. The proposed PCC algorithm was tested in Santa Maria Maior parish in Lisbon that has a complex ancient urban fabric. The GAMA agent-based platform was used in the two simulation studies. The first compared two scenarios (normal and COVID-19) in three major public spaces in Lisbon and the second focused on the simulation of a real-time tourism crowding stress analysis scenario of visitors' arrival at the Lisbon Cruise Terminal.

The results show the proposed algorithm's feasibility to determine the PCC of complex urban fabrics zones and its application as an initial reference value for the evaluation of real-time crowding stress, namely in simulations for assessing overtourism scenarios, both in public open spaces as in highly constrained urban fabrics.

Keywords: carrying capacity; agent-based modelling; tourism; overcrowding; overtourism

[This page has been intentionally left blank]

Contents

[This page has been intentionally left blank]

List of Figures

[This page has been intentionally left blank]

List of Tables

Acronyms

ABM	Agent-Based Modeling.
CC	Carrying Capacity.
ECC	Effective Carrying Capacity.
GIS	Geographic Information System.
INCD	Infraestrutura Nacional de Computação Distribuída.
LOS	Level Of Service.
OSM	OpenStreetMap.
PCC	Physical Carrying Capacity.
RCC	Real Carrying Capacity.
SFM	Social Force Model.
TCC	Tourism Carrying Capacity.
UNWTO	World Tourism Organization.
UTM	Universal Transpose Mercator.
VM	Virtual Machine.

Chapter 1.

Introduction

Contents

This chapter describes the motivation of this dissertation and presents the document's contributions and organization.

1.1 Motivation and research problem

Tourism has a major social and economic impact on many European countries, namely, Portugal. Up until 2019, before the start of the COVID-19 pandemic, tourist arrivals were steadily increasing in Portugal and across the world, as shown in Figures 1.2 and 1.1.

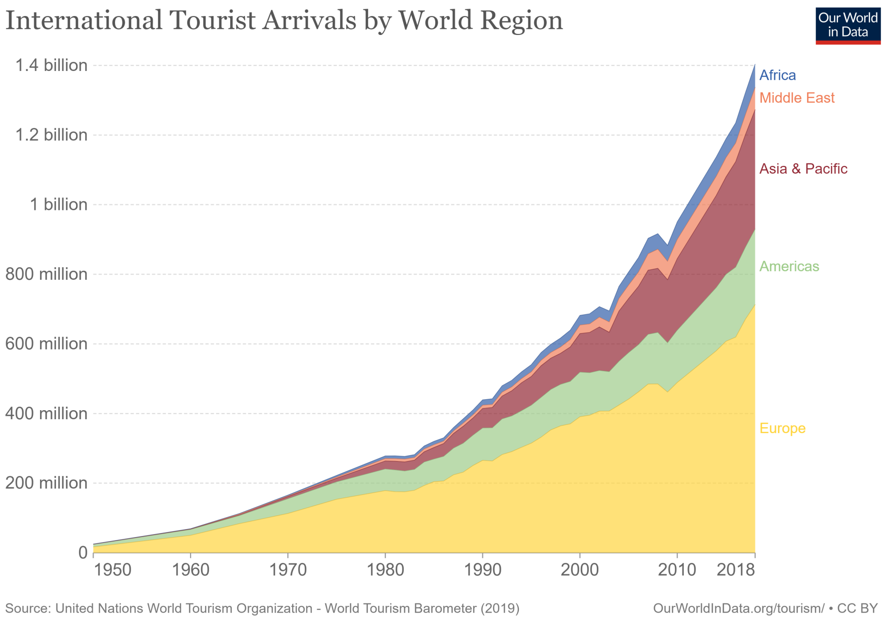

Figure 1.1: Number of international tourist arrivals by world region from 1950 to 2019

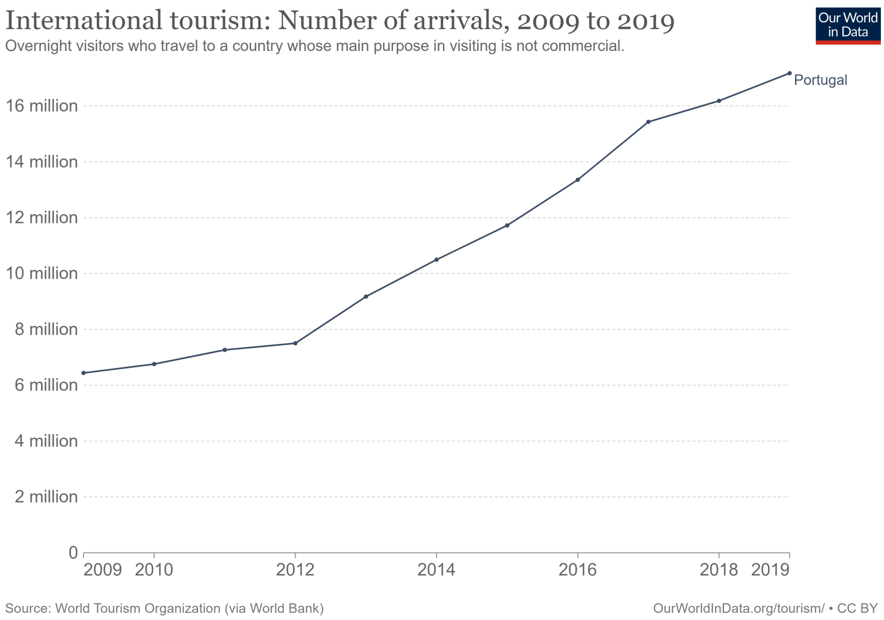

Figure 1.2: Number of tourist arrivals in Portugal from 2009 to 2019

The increase in tourism activity, although it contributes greatly to the economy, leads to the occurrence of what is known as *overtourism*. This phenomenon is defined by the World Tourism Organization (UNWTO) as *"The impact of tourism on a destination, or parts thereof, that* *excessively influences the perceived quality of life of citizens and/or quality of visitors' experiences in a negative way"* [69]. Furthermore, overtourism is considered detrimental to environment [28] and cultural heritage integrity [1].

In Figure 1.3, an example of a tourist destination being affected by overtourism can be observed.

Figure 1.3: Example of a touristic destination affected by overtourism[1]

Dealing with overtourism is mostly a matter of resources management (aka tourism governance [39]). Resources are by definition finite. Therefore, there should be a limit in their usage or consumption to avoid their depletion or degradation. Tourism offer itself is a finite resource. There are limits to how many tourists can be accommodated in a certain site or can visit a site simultaneously and still have a pleasant experience. The available usable area is limited, the tourism facilities and services are limited, and their levels of service are also limited, setting a general tourism usage limit that should not be surpassed to maintain the quality of tourism activities and the life of residents.

Overtourism's most obvious manifestation is a too high pedestrian density in visiting places. Dichter et. al [21] point out that one of the sub-elements for tourism strategy to cope with overtourism is determining the carrying capacity of a destination. The concept of Tourism Carrying Capacity (TCC) is an important concept at play and stands for the upper limit for the number of visitors in a given destination that allows maintaining the quality of tourism experience and avoiding the degradation of the environment [50]. Such threshold can be based on (i) a fine-grained measurement of the walkable area for visitors in the same destination, like the one proposed in this dissertation and, (ii) on Schumann's Social Distance Theory (1976), as suggested in [64] and also shared in this dissertation. To monitor in real-time if the TCC is being met, we require a real-time fine-grained measurement of the number of visitors in the destination, such as the one proposed in [20] or on mobile network usage data [4, 13].

[1] https://medium.com/lybra-tech/when-tourism-is-too-much-overtourism-causes-and-solutions-70d4f802f57c

It is also important to understand the dynamics of the interaction between the crowds and the environment, i.e., how the pedestrians (tourists) use the urban open space, and how the weather conditions tourism activity, to adapt the tourism control interventions to the reality of tourism in a destination. The set of interactions between these entities can be viewed as a complex system. Agent-based modeling and simulation can help understand complex systems, bring to the surface patterns and phenomena from simpler interactions between the intervening agents and the environment, mimicking the behavior of the system entities [42]. Tourism is considered a complex system, thus, it can benefit from this paradigm to understand and predict underlying phenomena [6]. To understand how the spatial configuration of urban fabrics and segments can handle the pedestrian activity, 2 agent-based models were developed in the scope of this dissertation, to simulate scenarios of egress from mass crowd events is public urban open spaces, and the simulation of tourist activity derived from cruise ship arrivals.

The UNWTO proposed a list of 7 strategies and a corresponding set of measures to address urban overtourism [69]. One is the deployment of localized tourism control interventions based on crowding state measurements. The PREVENT CROWDING project being developed at Iscte, where the research work of this dissertation is scoped, is an example of such an intervention, aiming at dispersing tourists through a mobile application that, based on real-time measurement and forecasting of crowding levels, suggests less visited and more sustainable route recommendations. This solution can help the fulfillment of some of the strategies pointed by the UNWTO [69], and is also framed in the eleventh UN sustainability goal: *"Make cities and human settlements inclusive, safe, resilient and sustainable"*, as represented in Figure 1.4. The PREVENT CROWDING project will be further described in Chapter 2.

Figure 1.4: United Nations Sustainability Development Goals[2]

1.2 Research questions and objectives

The work developed in the scope of this dissertation is explorative by addressing several overtourism concerns and scenarios. The concrete research questions that arose during our research were:

1. How to assess spatially fine-grained carrying capacity of a complex urban fabric?
2. How to model and simulate non-urgent egress from mass crowd events in open public spaces?

[2] https://sdgs.un.org/goals

3. How to model and simulate an urban fabric's capability to accommodate unusual arrivals of tourists?

To answer the previous research questions, the objectives to fulfill are the following:

- Develop an algorithm to create a polygonal representation of a complex urban fabric and estimate the pedestrian Carrying Capacity (CC) of segments, streets, and open spaces.
- Develop an agent-based model to simulate non-urgent egress from mass crowd events in 3 major Lisbon open spaces.
- Develop an agent-based model to simulate the urban fabric crowding stress, from the arrival of tourists from cruise ships, on the historical neighborhood of Alfama in Lisbon.

To fulfill the previous objectives, the following open source / open data technologies are used: the *OpenStreetMap*[3] worldwide geographic database, the *Python*[4] programming language and *GAMA*[5], a platform for building spatially explicit agent-based simulations.

1.3 Expected contributions

The main expected contributions of this dissertation are (1) algorithms for calculating the usable area for pedestrians (visitors) in a given urban space (e.g. street or square) and for the corresponding PCC; (2) an agent-based model for open space mass crowd event egress and, (3) an agent-based model to evaluate urban fabric crowding stress. The agent-based models aim at providing insights into two complementary situations which are typical in urban tourism overcrowding. The former corresponds to a scheduled / planned scenario in open spaces specifically chosen due to their high PCC, but where the crowding evolution must be known in advance to book the required resources (e.g. police, paramedics, urban cleaning teams). The latter corresponds to an organic / uncontrolled scenario where attractive visiting spots in constrained historical urban fabrics are often flooded by tourists. In both scenarios we use the aforementioned algorithms.

1.4 Dissertation organization

This dissertation is structured in 6 chapters. Chapter 2 presents the scope and important concepts of the dissertation's work. Each of the following 3 chapters describes a distinct study performed in the dissertation, including their research motivations, related work review, methods, and results: Chapter 3 presents the research on the algorithm for assessment of urban fabric CC, Chapter 4 presents the agent-based model of mass crowd urban open space events, and Chapter 5 presents the agent-based model for urban fabric crowding visitation evaluation. Finally, Chapter 6 contains concluding remarks for the 3 studies and presents directions for future research work.

[3] https://www.openstreetmap.org/
[4] https://www.python.org/
[5] https://gama-platform.org/

CHAPTER 2.

Scope and Concepts

Contents

In this chapter, we present the scope of the present work and key concepts addressed in this study

[This page has been intentionally left blank]

2.1 The PREVENT CROWDING project

The *PREVENT CROWDING* project, also known as *Sustainable Tourism Crowding* (STC) is a project in development in ISTAR (Information Sciences and Technologies and Architecture Research center) at the Iscte university. The main goal of this project is to mitigate overtourism, by offering a crowd-aware route recommendation system for tourists, which recommends routes for visiting less crowded, more sustainable points of interest for a given time period. A successful adhesion and usage of the system could help disperse tourism activity across an urban area, aiding the management of tourism offer. The system is aimed to be firstly implemented at the Santa Maria Maior district, one of the most affected zones in Lisbon by the overtourism phenomenon.

The goals of this project are aligned with UNWTO's strategy to address visitors' growth in cities: *"Promote the dispersal of visitors within the city and beyond"* [69]. The system uses a microservices architecture, a distributed computing architecture where the system is divided into multiple independent components, referred as services, which communicate with each other through their exposed interfaces [23].Each service has high cohesion, i.e., fulfils specific tasks and manage certain data domains, and low coupling, i.e., do not depend on the internal workings of the others. They are independently deployed, only needing to consume data from the other services, knowing their interfaces.

In Figure 2.1, the general overview of the system architecture is presented. Each cube represents a service, each circle represents an exposed interface of a service, and the arrows represent the dependencies between services. The diagram was drawn using AjiML, a graphical language for modelling microservice architectures [60].

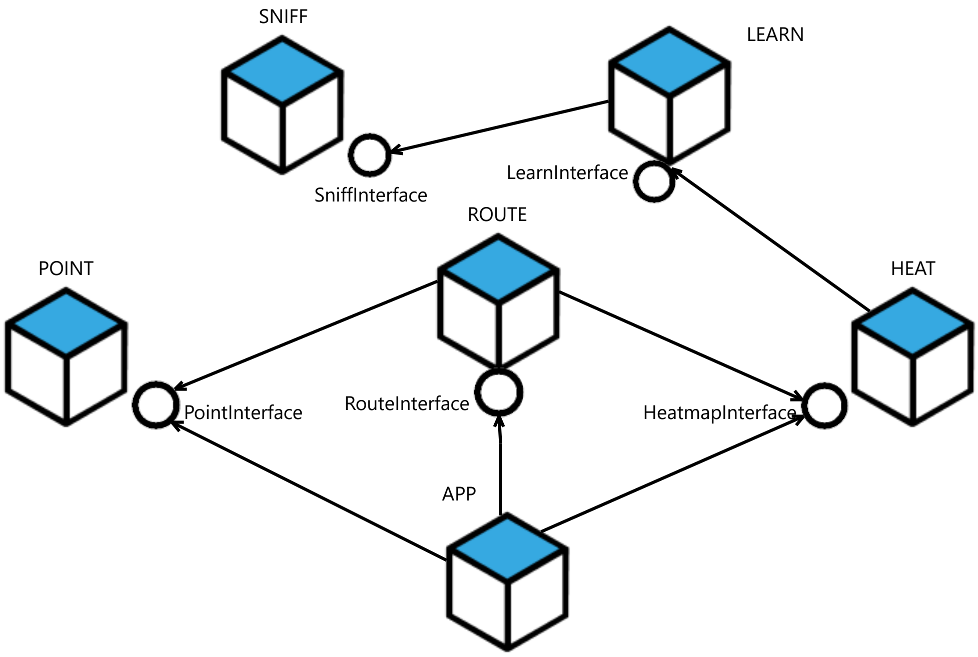

Figure 2.1: Overview diagram of PREVENT CROWDING system's architecture

A listing of the services integrating this system, along with a brief explanation of the roles and current advancements on each service follows:

- **SNIFF** – This service is represented by a network of smart sensors, which estimate crowd counts by counting nearby active mobile devices, tracking their multiple networks (Bluetooth, Wi-Fi, and Mobile Networks). Sensor prototypes built are described in [59] and were tested in outdoor and indoor scenarios ;

- **LEARN** – This service processes and stores all data from the *SNIFF* service, using Machine Learning algorithms to generate spatial-temporal forecasts of crowding patterns in the covered zones. The forecast data is then served by the service, being consumed by the *HEAT* service.

- **HEAT** – This service, consuming real-time occupancy and forecasting data from *LEARN*, and receiving location pings from mobile devices with the application (*APP*), creates crowding contour lines.

- **POINT** – This service hosts and serves the information about the georeferenced points of interest of the covered zones. It can be edited via a web platform interface by local municipalities and other stakeholders. A prototype of this service is presented in [51].

- **ROUTE** – This is a central service that provides crowd-avoiding, sustainability-maximising route recommendations, based on given parameters and preferences. It uses point of interest data from *POINT*, the crowding contour line data from *HEAT*, and information provided by the mobile application (*APP*). A first implementation of the routing recommendation algorithm is presented in [3].

- **APP** – A mobile application that records tourists' preferences and constraints and provides them to the *ROUTE* system which generates the route recommendation that is shown to the user in the interface. Also functions as a monitoring tool that feeds useful information (e.g. user's location) to the system. A first version of this application is described in [40].

The work presented in this dissertation spans multiple services. The CC assessment of urban fabrics is important for the services covering detection and quantification of overcrowding, and the agent-based simulations give insight on the effects of urban fabrics on the fruition of public space by a large number of visitors, which can aid the correct placement of sensors and learning mechanisms for crowding prediction.

2.2 The OpenStreetMap initiative

OSM[1] is a worldwide open data geographic mapping project developed, since 2004, through the contribution of a large community of volunteers spread across the globe. The mapping elements are stored in a geographic database which can be used without limitations, as long as credit is given. Currently OSM has a wide coverage of the earth surface and major cities are represented with significant accuracy.

[1] https://www.openstreetmap.org/about

2.2.1 OpenStreetMap data

The data gathered, stored and served by the project consists on an immense set of geographic entities, referenced in OSM's data model as elements. These elements are categorized by three main types:

- **Node**: Represents a specific point on the surface of the earth. It is defined by its coordinates (latitude and longitude) and a node identifier. Nodes can be used to define standalone point features, such as points of interest, street elements (e.g. benches, trees, statues) or simply define points contained on a way or relation.

- **Way**: A way represents a multi-point geometry, consisting on an ordered list of nodes, with a minimum of 2 nodes, and a maximum of 2,000 nodes. Usually used to define linear or multi-linear elements (e.g., roads, streets, water streams, public transport paths), and areas (e.g., buildings, squares, gardens, parks).

- **Relation**: Represents an ordered list of nodes, ways or relations, defining a relationship between its members. The geographic entities modeled by a relation include multi-polygons (group of polygons considered a single entity), administrative boundaries, routes, etc.

Elements are characterized by tags. Tags consist on key-value pairs used to describe specific features of elements. The key describes a topic, category, or type of feature, while the value provides detail for the key-specified feature. The tagging system follows an agreed convention to allow for accurate characterization of elements. In Figure 2.2, the element representing the central pavement of the Praça da Figueira square is selected in the map of the main OSM website [2], showing to the left the tags and respective values of the element.

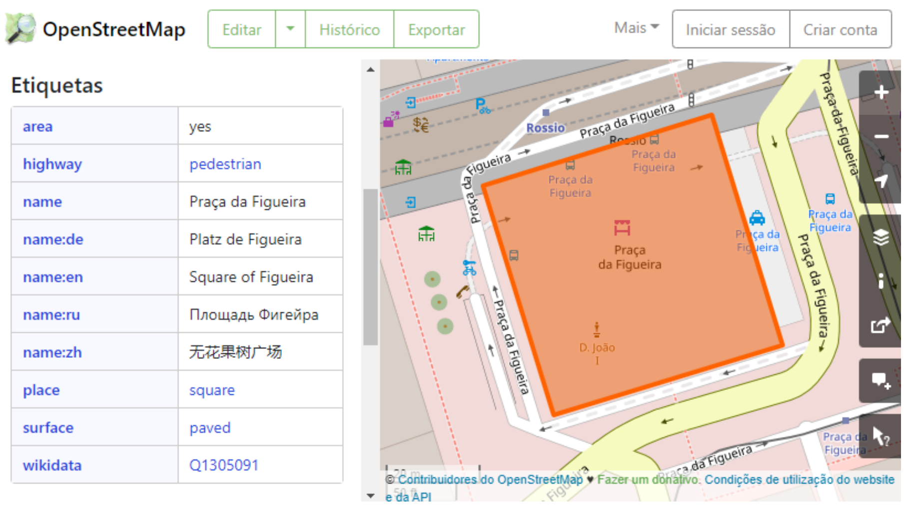

Figure 2.2: Screenshot of the main OSM page, showing Praça da Figueira in the map and its tag-value mapping

[2] https://www.openstreetmap.org/

2.2.2 Data access

Data can be accessed mainly via the following sources:

- **Planet.osm file**: A file containing all existing information on the OSM database. Can be downloaded via torrents or mirrors[3].
- **Export files**: Files obtained through the OSM's main website, which allow to select all elements within a latitude-longitude bounding box[4].
- **Overpass API**: An Application Programming Interface (API) which allows for custom queries to obtain specific types of elements [5].
- **Third-party providers**: Entities, external to the mapping project, which provide raw, derived or pre-processed OSM data.

Files can be formatted in plain XML, which can be read with any common XML parsing tool, or PBF, a specialized binary format file which occupies less space, allows for faster parsing, but requires specialized tools to read.

2.3 Carrying capacity

CC, originally from the field of ecology, is defined as the maximum supported population number for a given environment, until there are no sufficient resources to maintain population growth, and degradation starts to occur. One of the first usages of this concept in the Malthusian's population theory [58], which states the exponential growth of human population is limited to existing resources (which cannot grow exponentially), turning the growth model into a logistic model, being its upper limit the CC of the human population. The concept has been applied in multiple other fields, such as economy, biology, and population studies [38], and has had various definition iterations according to its application.

The concept has been approached in tourism studies, being known as TCC. According to the World Tourism Organization (WTO), it is defined as *"the maximum number of people that may visit a tourist destination at the same time, without causing destruction of the physical, economic, social-cultural environment and an unacceptable decrease in the quality of visitors' satisfaction"* [50]. Getz [24] stated there are 6 TCC categories: physical, economic, perceptive, social, ecological, and political. Respectively, each category represents the maximum usage of a touristic resource before the occurrence of physical deterioration, excessive economical dependence, perception of overcrowding by tourists, ecological damage to natural and political instability.

There is not a consensus in the utility of numerical values for TCC. McCool et al. [45] assert that these values are "magic values", being dependent on a large number of realistically hard to model variables, and which can lack stability, and inappropriate to solve the problems of tourism development. Nonetheless, these authors acknowledge the usefulness of previous TCC research for managing tourist sites.

[3] https://planet.openstreetmap.org/
[4] https://www.openstreetmap.org/export
[5] https://dev.overpass-api.de/overpass-doc/en/

Although TCC is considered an imperfect measure for the actual CC of a touristic place of interest, it can be used as a reference value to measure overcrowding and compare. As such, in this dissertation, we use the concept of TCC as a practical threshold for detecting tourism overcrowding in the scenarios of the *PREVENT CROWDING* project.

Cifuentes presented a methodology, widely used in the literature, which is capable of estimating values or ranges of TCC in different types of touristic sites [16]. The researcher defines 3 levels of TCC:

1. **PCC**: This level of TCC represents the maximum number of visitors that can be physically accomodated into an area over a time period. It is calculated using the following formula:

$$PCC = \frac{A}{Ap} \times Rf$$

where A is the available area for tourism activity, Ap is the area used per tourist for a specific type of tourism activity, and Rf is the rotation factor, which represents the number of visits within a period of time (usually a day). To assess instant PCC, the rotation factor can be omitted from the equation. The Ap parameter is culture and activity dependent, as different cultures perceive the ideal space per pedestrian differently, and each tourism activity requires different spaces per pedestrian.

2. **Real Carrying Capacity (RCC)**: This value is derived from a previously calculated PCC, applying corrective factors with different natures (e.g., physical, ecological, economical), specific to the location. The formula to assess RCC is:

$$RCC = PCC \times (cf1 \times cf2 \times \ldots cfn)$$

where *cf1...cfn* are the defined corrective factors. Each factor is calculated using the expression:

$$cf = 1 - \frac{Lm}{Tm}$$

where Lm is the limiting magnitude of the factor, and Tm is the total magnitude of the factor.

3. **Effective Carrying Capacity (ECC)**: The final level of TCC is devired from the RCC, multiplying by

$$ECC = RCC \times Mc$$

where Mc is the management capacity of the site. This variable is often determined from the adequacy of the available infrastructure, equipment and staff for the tourism activity.

Obtaining the final stage of TCC (ECC) requires further research on the significant corrective factors and management capacity of an urban tourism site. As such, assessing the levels of RCC and ECC are not on the scope of this dissertation. Only the PCC will be considered.

2.4 Agent-based modelling

The known world moves through a vast set of complex systems, each functioning through the interaction of multiple entities. In order to gain insight of said systems, often times it is impractical or impossible to directly observe certain real or hypothetical scenarios, and can be hard to mathematically or statistically model. Agent-Based Modeling (ABM), also known as *multi-agent* or *individual-based* modelling, can be successfully used to study these systems. It is generally described as the computational modeling of a real-life system as a collection of autonomous entities, defined as agents, which behave and take decisions based on a set of rules, and can interact with each other and the modeled environment [8]. This model allows to simulate a system, based on a set of defined variables, allowing the discovery of emergent dynamics from the complex set of interactions and entities. Macal [42] recognizes the difficulty of the acceptance of a single definition of the concept, and proposes 4 different definitions of ABM, based on individuality (by definition, every model will have individual, heterogeneous agents), behavior, interactions and adaptability:

- **Individual ABM**: Agents have individual and diverse characteristics, but behave by external scripting, not acting on internal state. Interactions are limited. Agents are not adaptive.

- **Autonomous ABM**: An individual ABM where agents will have autonomous behavior, evaluating external conditions and internal state. Interactions are limited. Agents are not adaptive.

- **Interactive ABM**: An autonomous ABM where agents interact with each other and the environment. Agents are not adaptive.

- **Adaptive ABM**: An interactive ABM where agents learn from their "experience"and adapt their behaviors throughout a simulation.

A typical ABM has the following procedure: In each simulation cycle, each agent evaluates the current situation of the modeled environment, through a set of defined environment variables and other state indicators, which can be altered by the behaviors of other agents (including the self), and the internal state caused by the interactions of the other agents with themselves. Depending on the set behaviors and rules of the agent, a set of actions could be triggered, possibly affecting the state of the surrounding entities. Environments can also be defined as agents, having their set of procedures to change state accordingly to the system's situation. ABM can be applied to almost any domain. There are uses of this model in a wide variety of domains, such as physics, biology, ecology, geography, tourism, psychology, etc. [42]

The concept of *digital twinning* is also tied to ABM. It consists on the modeling of digital counterparts of real-life entities with some level of detail and fidelity, fed by dynamic real data [43]. In the ABM presented in this dissertation, to some degree we can affirm that the modeled environments are effectively static digital twins of their real-life counterparts, in terms of geometry, which is the only attributes described in the models. The *DIGITALTWINS* project, led by Prof. Dr. Inês Boavida-Portugal, aims to harness digital twinning to mimic multi-dimension dynamics of tourism destinations, integrating stakeholders' inputs to help lead the tourism complex adaptive system towards a sustainable future.

In this dissertation, we proceed on modeling visitor open space egress and urban PCC stress derived from worst-case scenario of tourist arrival using this paradigm, studying the outcome of a series of stochastic simulations with different scenarios.

2.5 Social Force Model for pedestrian dynamics

There is a general perception that the human behavior is irregular and unpredictable [30], but in relatively simple situations, the individual behavior of a pedestrian could be modeled. Lewin [41] presented the idea that behavior changes are guided by *social forces*, represented by the sensory stimulus caused by the situation and environment, which cause a behavioral reaction, depending on a person's personal aims and interests.

The Social Force Model (SFM), presented by Helbing and Molnar [30], is a microscopic mathematical model that represents pedestrian dynamics and interactions, based on the concept of *social forces*, which collectively act upon the decision-making of the movement of pedestrians, resulting in the manifestation of a decided velocity. There are 3 main terms regarding the assessment of the driving force of the actual velocity of a pedestrian:

- The force towards the desired velocity, dependent on the desired walking speed and the directions towards the destination;

- Repulsive forces from other pedestrians and obstacles (e.g., borders, walls, statues), to avoid and keep distance from them;

- The attractive forces from other possible sources (e.g., a shopping window on a way to a destination). In practical applications of SFM, this term is usually neglected.

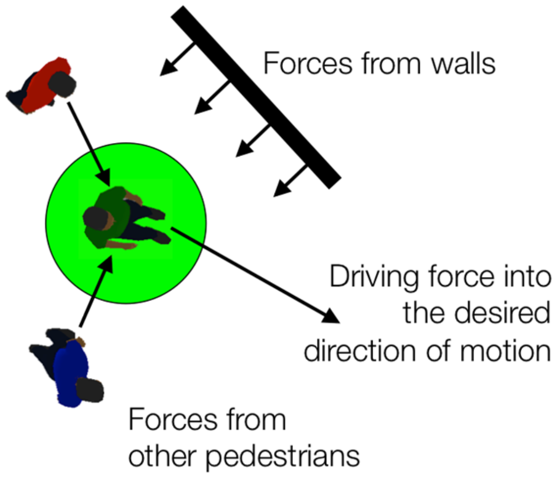

Figure 2.3: Graphic representation of the Social Force Model (source: [29])

The SFM can be naturally implemented in an agent-based simulation, as each pedestrian can be represented by an autonomous agent, which interact with each other by applying the social forces, and decide which velocity to take based of the sum of all forces.

2.6 GAMA platform

The GAMA platform [6] is an Eclipse-based [7] Integrated Development Environment (IDE) for spatially explicit, agent-based and agent-oriented modeling and simulation [63]. The platform is designed to develop and run complex, large-scale models, with a wide support for integrating and visualizing Geographic Information System (GIS) data.

Models are written with the GAML language, an agent-oriented language (i.e, a language where almost every active entity is considered an agent), with a rich set of simple and complex operations on every supported data type. The language is intuitive to computer scientists and programmers, but accessible to scientists without computer science backgrounds, allowing everyone to grasp modeling with this language. It also allows the usage of multiple paradigms

Each model can have one or more experiments, which represent simulation scenarios, where model inputs, variables and outputs can be defined. Experiments act as agents that control the execution of the model. Agents are defined through *species*, which are analogous to classes in object-oriented programming. Species contain the definition of agent attributes and behaviors. Actions of an agent that trigger at each simulation cycle, given the pre-conditions are met, are called *reflexes*. These actions are used to model the behaviors of agents. Species can inherit attributes and behaviors from parent species, and from *skills*, which are built-in modules containing presets of attributes and behaviors to give the agents a role in a specific domain (e.g., driving skill, pedestrian skill, network skill).

GAMA comes with a *headless* mode. This mode allows to run resource-intensive simulations in the background, without an user interface, allocating the needed resources for each. It is useful, e.g., to run heavy simulations in resourceful remote machines. With this mode, the simulation outputs can be stored on the machine in various formats for later analysis.

On Figure 2.4, the GAMA platform user interface is displayed, showing the execution of a Predator-Prey example model, input variables and outputs.

The version used in the development and simulation of the models on this dissertation is the GAMA 1.8.2 alpha build, which comes with the built-in Pedestrian skill, which allows to model pedestrian walking behavior according to the SFM proposed by Helbing and Molnár [30], allowing to define values for its parameters, or use a simplified version with pre-defined values.

2.7 Rapid systematic literature review

For each individual study presented in this dissertation, a literature review is conducted. Reviewing the related work allows us to better understand each field of study, gives us insight of the state of the art of the subjects, discovering what others have done to solve common

[6] https://gama-platform.org/
[7] https://www.eclipse.org/

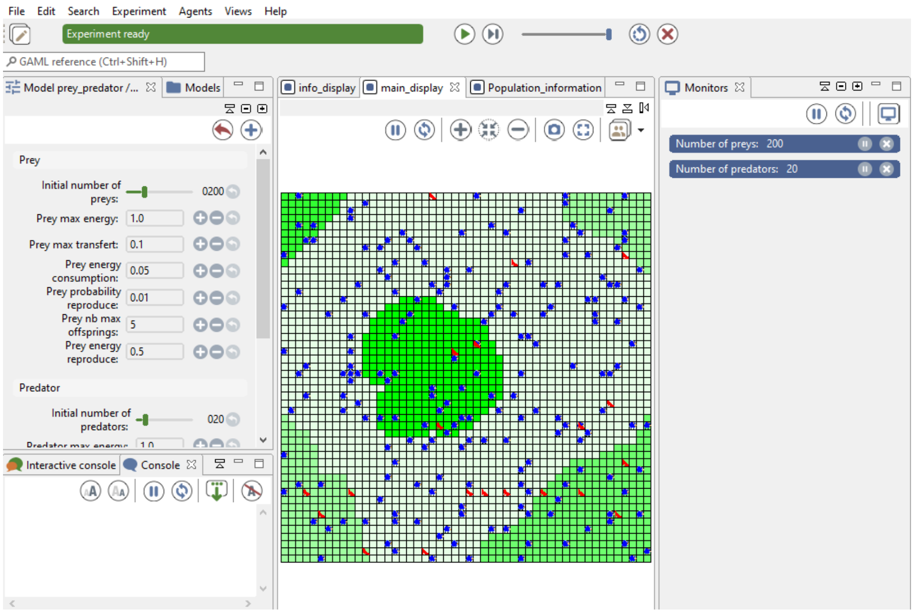

Figure 2.4: GAMA Platform UI, showing a running simulation

problems, what approaches were taken and the application of the studies. The reviews were accomplished by using a Rapid Review (RR) methodology [14], which enables for a structured scientific synthesis of current knowledge in a short period of time [66]. The literature review studies will consist on the gathering, filtering, organizing, analysis and categorization of the relevant related studies. The reviews will follow the next structure:

1. **Research objectives**: Highlighting of the main objectives of the review research;
2. **Review methodology**:
   a) **Study selection criteria**: A brief presentation of the criteria of study selection.
   b) **Inclusion and exclusion criteria**: Shows tables with the enumeration of the inclusion and exclusion criteria for study selection;
   c) **Search strategy**: Explaining the study search strategy, referring the used scientific document sources and other study gathering techniques (e.g., snowballing);
   d) **Search strings**: Present the defined combinations of search keywords to find studies in the scientific databases and search engines;
   e) **Selection process**: Description and diagram of the entire study selection process;
   f) **Validity threats**: Highlight of the possible threats to the validity of the review study.
3. **Classification taxonomy**: This section presents the analysis of the found related work. A taxonomy is proposed to categorize the studies by their major aspects, showing their similarities and differences, how they relate to each other and their strengths and weaknesses. This is an objective and methodical framework for article reviewing, reducing ambiguity and increasing rigor on the survey. Each category and possible values are explained.
4. **Related work review**: In this section, each selected studies are analysed and categorized. For each related study, the following analysis will be presented:

- **Objective**: Motive for the study.
- **Technical summary**: A summary of how the approaches presented in the studies work. Includes advantages, disadvantages and expected future work noted by the authors.
- **Classification**: A table containing the taxonomy attributed to the study.

A summary table with the classification of all analysed studies is presented at the end.

5. **Conclusions**: The main conclusions obtained from the review are presented.

Chapter 3.

Carrying Capacity

Contents

This chapter presents the motivation, related work and the description of the algorithm used to assess physical carrying capacity of an urban fabric, as well as results and validation.

3.1 Motivation

As previously stated in this dissertation, when the CC of a resource is reached, social, economic and environmental damage may occur, affecting the tourism experience and the quality of life of the residents. Cities across the world are victims of tourism-induced overcrowding, due to the steadily growing tourism activity. This overcrowding can be viewed as the surpassing of the CC of a location.

Assessing the CC of open space urban places of interest is important in effective public space usage management, allowing for cities to better allocate resources (e.g., security forces, cleaning teams, emergency personnel) and devise adequate measures for tourism control. Real-time monitoring of the state of the touristic usage of a place of interest is crucial to apply real-time tourist control responses, such as the proposed system currently being developed in the PREVENT CROWDING project. For an effective real-time measurement of the overcrowding of places of interest, reference values are needed, to objectively classify and compare crowding levels. The PCC is an adequate initial measure of overcrowding threshold. A fine-grained zoning of the available area for tourism use is also needed for adequate spatial tourist balancing strategies, promoting the visitation of less visited places and reduce overcrowding in usually crowded places. The area segmentation and subsequent assessment of the CC of a wide complex urban fabric can be challenging when there is a lack of official accurate spatial data about the open area of cities. Manually measuring every open space in this type and dimension of site also not feasible in a short time period, and requires some amount of resources.

The objective of the study of this chapter is to develop an algorithm to estimate instant PCC for open space zones of a complex urban fabric, using a Python script for spatial zoning determination, using OSM geographic open data, and apply it on the *Santa Maria Maior* district, Lisbon.

3.2 Related work

In this section, we present the analysis of relevant work to assess the PCC of tourism areas, following the previously explained Rapid Review methodology.

3.2.1 Research objectives

The objectives of this literature review are:

- Undertake rapid systematic reviews of research on PCC assessment of tourist places of interest;
- Filter the articles by relevancy to the given topics to further review in more detail;
- Integrate information from these studies;
- Determine how the various topics approached in the reviews relate to each other through taxonomic classification;
- Identify opportunities for improvement/further investigation.

Concretely, we want to determine:

1. What is the current body of knowledge about the approaches on PCC assessment of touristic places of interest?

2. What are currently the most effective approaches used for PCC assessment?

3.2.2 Review methodology

3.2.2.1 Study selection criteria

Selected articles must present empirical data or the production of publicly available, usable and testable artifacts. The inclusion of studies will not be restricted to any specific type of intervention, but real case scenarios are preferred.

3.2.2.2 Inclusion and exclusion criteria

To include or exclude found articles in this review, the defined criteria are in Tables 3.1 and 3.2.

Table 3.1: Inclusion criteria

Inclusion criterion	Description
IC1	Studies must be published in journals after 2011
IC2	Articles must be related to the assessment of PCC of touristic places of interest
IC3	Articles must be research-based
IC4	Studies must be peer-reviewed

Table 3.2: Exclusion criteria

Exclusion criterion	Description
EC1	Articles are inaccessible
EC2	Articles are not written in English

3.2.2.3 Search strategy

The adopted strategy was to search related work on *Scopus*, a scientific electronic bibliographic database with a wide coverage of disciplines, including Computer Science and Tourism related studies. To reach more related documents and lessen the chance of neglecting useful related work, we performed the snowballing procedure on the relevant articles found via the search engine. Snowballing is a known technique to identify related articles based on citations [61]. Forward snowballing consists of searching which articles referenced the current one, while backward snowballing is searching for articles referenced by the starting article.

3.2.2.4 Search strings

Search strings were created through trial and error, combining multiple terms related to the domains of each study. To build a search string for the *Scopus* database, the alternative terms of a domain are joined with an *OR* logic operator, indicating that any of the terms can appear. Domain term groups are then joined with the *AND* operator, asserting the presence of each domain on the search results. When we want to exclude a specific term from the search results, a *NOT* operator is inserted before the term. A domain comprised of multiple words must be cased with quotes. The search strings used to search on the database are the following:

- (pedestrian OR visitor) AND "physical carrying capacity"AND (estimation OR rating OR calculation OR determination OR evaluation OR assessment)

3.2.2.5 Selection process

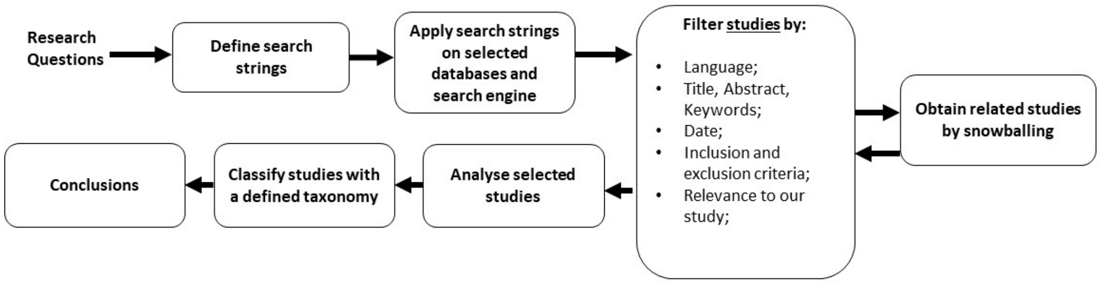

Figure 3.1: Diagram of the articles selection process

The selection process steps are defined as:

1. Execute the search string on the defined databases and search engines;
2. Apply the inclusion criteria (orderly by title, abstract, keywords and content);
3. Apply the exclusion criteria;
4. Apply backward and forward snowballing to the remaining articles;
5. Repeat steps 2 and 3 on the new articles selected by snowballing.

In Figure 3.1, a diagram of the selection process can be visualized.

3.2.2.6 Validity threats

This review has possible threats to its validity. First, the number of articles used for the analysis of the related work may be less than ideal to assure the validity of the results. Out of all the related work found, only a total of 15 articles were selected.

First, only using a single search engine may affect validity, as there may be useful related studies which have not been properly indexed, missing on them.

Second, more variations of the search strings, such as appending more terms and varying term combinations, could have been used to narrow the database results to a lesser number of more relevant articles. Limiting the search strings might also have led to not finding some related work.

Last, inaccessible articles were discarded, which could lead to the loss of important information. An effort to reach the authors should have been done.

3.2.3 Classification taxonomy

3.2.3.1 Taxonomy criteria

To categorize the different types of TCC assessment studies of tourism places of interest, a set of criteria was defined. The Table 3.3 shows the selected criteria.

Table 3.3: Taxonomy criteria

Criterion	Acronym
Tourist destination type	DT
Pandemic Awareness	PA
Spatial analysis	SA
Temporal analysis	TA
Validation	V

3.2.3.2 Tourist destination type (DT)

The studies on CC assessment can be categorized by the type of tourism activity and environment. This criterion lists the site tourism categories, being defined on a nominal scale with the following categories:

- **Religious Site**(RS): The site is a religious place of interest, a destination for mass pilgrimage and visiting during sacred seasons.
- **Beach**(B): The site to assess is an oceanic or fluvial beach, with a typical activity of bathing, sunbathing, water and sand sports.
- **Natural Park**(NP): The tourism site is a natural park, a protected area with a typical usage for ecotourism, allowing for a great contact with nature.
- **Archaeological Site**(AS): The place has archaeological tourism interest, with ancient monuments, relics and structures.
- **Urban space**(US): The tourism site is an urban environment (e.g., historical neighbourhood, city squares, urban monuments).

3.2.3.3 Pandemic Awareness (PA)

This criterion describes if the study uses pandemic-related social distancing guideline for the assessment of TCC. The ordinal scale representing the criterion is the following:

1. **Not-Aware**: The studies do not take into consideration the pandemic scenarios of social distancing.
2. **Aware**: Studies do take into consideration the pandemic-forced social distancing.

3.2.3.4 Spatial analysis (SA)

The CC assessment studies have different approaches on the spatial grain of the tourism place of interest. This criterion describes the spatial analysis grain of the studies, being represented in an ordinal scale by the following values:

1. **Total Site Area**: The studies do not differentiate specific zones of the tourism site, applying the same methodology across the site.

2. **Zoning**: Section a site into smaller zones, emphasizing the different tourism usages of different places in the same area.

3.2.3.5 Temporal analysis(TA)

CC can be analysed in time, from the CC of an instant (maximum number of people at a given instant of time), to the CC of a place in a day, month, year, season, etc. The criterion lists the temporal analysis grain of the CC, and is represented by a nominal scale:

- **Period**(P): CC is analysed for a time interval (during a day or more).

- **Instant**(I): CC is analysed for an instant of time.

3.2.3.6 Validation (V)

This criterion states if the CC assessment is validated or not.The following ordinal scale categorizes the criterion:

1. **Not Validated**: The study is not validated.

2. **Validated**: The study is validated.

3.2.4 Related work review

3.2.4.1 Zacarias et al.,*"Recreation carrying capacity estimations to support beach management at Praia de Faro, Portugal"*, 2011 [71]

Objective: The main objective of the present study is to estimate physical-ecological and social-cultural CC of the Praia de Faro beach, Portugal.

Technical Summary: To estimate the physical-ecological CC of the beaches of Faro, Zacarias et al. apply the Cifuentes methodology [16]. The PCC was determined for 2 scenarios of visitor density: 5 m2, and 10 m2 per person. The variables for the assessment were obtained via surveys and in-field observation. The social-cultural CC was assessed through statistical analysis of questionnaire answers during a survey on the beach.

As future work, Zacarias et al. note that applying other frameworks based on upper limits of acceptance is needed to identify ´the sustainable beach load and codes of conduct for beach users.

Classification:

Table 3.4: Classification of *"Recreation carrying capacity estimations to support beach management at Praia de Faro, Portugal"*

Source	DT	PA	SA	TA	V
[71]	B	1	1	P	1

3.2.4.2 Göktuğ et al.,*"Carrying capacity assessment of Tortum Waterfall, Turkey"*,2013 [26]

Objective: The purpose of the study was to estimate the pedestrian carrying capacity of the recreation area around the Tortum Waterfall, in Turkey, where the increasing number of visits induces stress to the walking facilities and decrease visitor satisfaction.

Technical Summary: The authors used the method of Highway Capacity Manual (HCM) [54], adapted to pedestrian walkways and viewing platforms by Itami et al. [36].

For National Park walkways and viewing platforms, the selected acceptable LOS was Level D, and for stairs, the level C was selected as adequate. Using the method, the CC of the viewing platform, the stairs and walkways was calculated. The total CC of the route is then assessed.

The study then assesses the adequacy of the pedestrian facilities' widths, concluding that the width at the time of the study (1.2 m) was not appropriate, pointing a required width of 1.5 meters.

The authors do not point possible future work, but advise on the reform of the waterfall facilities, increasing the visitor capacity.

Classification:

Table 3.5: Classification of *"Carrying capacity assessment of Tortum Waterfall, Turkey"*

Source	DT	PA	SA	TA	V
[26]	NP	1	2	I	1

3.2.4.3 Göktuğ et al.,*"Estimating carrying capacity of Olympos Bey Mountains Coastal National Park"*,2013 [25]

:

Objective: The study serves the purpose of determining the carrying capacity of Olympos Bey Mountains Coastal National Park, to allow for its adequate use and management.

Technical Summary: The authors used the methodology of Cifuentes [16] to assess the ECC of 5 bays of the National Park. The corrective factors and variables for the assessment were determined via in-field observations and surveys.

The authors conclude that the PCC and RCC could overall accommodate the recorded visits, depending of the time of the year, but due to the lack of staff, the ECC of 690 visitors/day for the sum of the bays could not accommodate the visiting rate assessed during the time of study.

As future work, it is suggested the undertaking of social CC and visitor satisfaction studies, and suggested to increase staff personnel.

Classification:

Table 3.6: Classification of *"Estimating carrying capacity of Olympos Bey Mountains Coastal National Park"*

Source	DT	PA	SA	TA	V
[25]	NP;B	1	2	P	1

3.2.4.4 De Sousa et al.,*"Tourism carrying capacity on estuarine beaches in the Brazilian Amazon region"*,2014 [19]

:

Objective: The present study investigates the estimation of the tourism carrying capacity of three touristic Brazilian Amazon beaches, for the future planning of sustainable development of recreational activities and the conservation of natural resources.

Technical Summary: The study investigates 3 beaches from Brazillian Amazon Coast: Colares, Marudá and Murubira.

To assess the CC of the beaches, it is used the Cifuentes method [16], but without the time parameter, following the adapted approach of Ruschmann et al. [55] and Zacarias et al [71].

The calculation of the PCC used the available beach areas, the ideal available space per pedestrian, determined as 10 m2 per person, defined by Ruschmann et al. The remaining variables for CC assessment were obtained via surveys and in-field observation. The ECC assessment was not only determined by beach, but by usage zone type and time of the day. Each beach was divided into 3 zones, based on recreational use patterns and topographic characteristics.

Comparing the observed visitor densities with the assessed ECC of each beach, the authors concluded the tolerable visiting limits were, overall, exceeded.

It is not suggested future work for this study, but coastal management measures are suggested to improve tourist activities.

Classification:

Table 3.7: Classification of *"Tourism carrying capacity on estuarine beaches in the Brazilian Amazon region"*

Source	DT	PA	SA	TA	V
[19]	B	1	2	P	1

3.2.4.5 Cisneros et al.,*"Beach carrying capacity assessment through image processing tools for coastal management"*,2016 [33]

Objective: The objective of this study is to present a methodology to estimate the Beach Carrying Capacity (BCC) and real-time beach usage level in coastal cities, using on-site information and image processing.

Technical Summary: The methodology of Cifuentes [16] was used to estimate the CC.

For the usable area of the beach, it is considered the worst case scenario of high tide, where there is the least usable beach area.

The PCC is calculated in 3 occupancy scenarios, each with a different pedestrian density. These scenarios are based on a model defined by Norma-Cubana [49], and represent low, medium and high occupancy, with, respectively, 25, 10 and 5 m2 of space per visitor. The remaining variables for the estimation of CC are obtained through in-field observations and surveys.

The authors also developed a visitor counting system based on camera image processing. The CC assessment method and counting system are applied and tested in Monte Hermoso, Argentina, dividing the beach into 6 zones. An analysis of socio-cultural factors that influence the CC of Monte Hermoso was also done, characterizing the touristic profile of the beach during summer.

It is concluded that the methodology and system can be used successfully to assess BCC, monitor beach activity and characterize usage patterns. The authors state that future work was being done to add and improve video analysis features to the system.

Classification:

Table 3.8: Classification of *"Beach carrying capacity assessment through image processing tools for coastal management"*

Source	DT	PA	SA	TA	V
[33]	B	1	1	P	1

3.2.4.6 Jangra et al.,*"Assessment of physical carrying capacity for managing sustainability at religious tourist destinations"*,2017 [37]

Objective:The present study focuses on developing a method for assessing the PCC of Braham Sarovar, a religious lake in Kurukshetra, Haryana, with an immense number of visit on specific occasions. With the PCC, it is pretended to evaluate of tourism sustainability of the place.

Technical Summary: The PCC is estimated using the Cifuentes method [16]. It is calculated for each zone, using the available area of each, assessed through *Quickbird* imagery from *Google Earth*.

The ideal available space per pedestrian and rotation factor are determined for each zone, through field survey and interviews with 50 pilgrims. The RCC and ECC were not calculated in this study.

The PCC of the religious site is complemented with parking and footway capacity, backed by the Indian Road Congress (IRC) standards, drinking water capacity calculated for certain levels of service, and toilet and bathroom capacity according to Government of India norms.

It is concluded that basic facilities are sufficient for normal days, but makeshift temporary facilities should be installed for peak days of visiting, such as days of sun eclipses.

The authors do not suggest future research work.

Classification:

Table 3.9: Classification of *"Assessment of physical carrying capacity for managing sustainability at religious tourist destinations"*

Source	DT	PA	SA	TA	V
[37]	RS	1	2	P	1

3.2.4.7 Rodella et al.,*"Assessment of the relationship between geomorphological evolution, carrying capacity and users' perception: Case studies in Emilia-Romagna (Italy)"*,2017 [53]

Objective: This objective of this study is to analyze the geomorphological characteristics, the users' perception and beach carrying capacity of two coastal stretches in Emilia-Romagna.

Technical Summary: In the present study, the CC is assessed through two different methodologies: The Cifuentes method [16] and the Williams and Micallef [46] method, a method for beach CC assessment. The authors suggest the use of the former is more adequate when user perception data is available, as opposed to the latter, which should be used when tourist preferences are not known. The variables for CC assessment with both methodologies were obtained via aerial images, in-field observation and questionnaires.

It is concluded that the two CC methodologies, applying the right corrective factors, give coherent results in comparison with the users' reference number. No possible future work is presented.

Classification:

Table 3.10: Classification of *"Assessment of the relationship between geomorphological evolution, carrying capacity and users' perception: Case studies in Emilia-Romagna (Italy)"*

Source	DT	PA	SA	TA	V
[53]	B	1	2	P	2

3.2.4.8 Sadikin et al.,*"Carrying capacity to preserve biodiversity on ecotourism in Mount Rinjani National Park, Indonesia"*,2017 [56]

Objective: This study aimed to analyze the land suitability for ecotourism based on specific criteria and the CC of the ecotourism area Mount Rinjani National Park, Indonesia.

Technical Summary: The CC is assessed for each different carried out activity in the National Park (sightseeing, relaxing, bird watching, orchid observation, camping, fishing, and bathing), all activities and all camping activities.

The CC was assessed in the present study using an equation defined by Douglass [22] for forest recreation CC.

This equation uses the following parameters: area required per activity, total demand per activity, area per person, capacity days per year and turnover factor.

The authors also perform a land suitability study on the area. Accounting the actual visitors per year in 2014 of 44.112 for camping activity, Sadikin et al. concluded the camping CC of the park was being exceeded. The results of the land suitability study shown ecotourism in the park was applied on moderately and marginally suitable land. It is not suggested future work.

Classification:

Table 3.11: Classification of *"Carrying capacity to preserve biodiversity on ecotourism in Mount Rinjani National Park, Indonesia"*

Source	DT	PA	SA	TA	V
[56]	NP	1	1	P	1

3.2.4.9 Morales et al.,*"Integrated Assessment of Recreational Quality and Carrying Capacity of an Urban Beach"*, 2018 [47]

Objective: The objective of this study was the assessment of the overall beach quality of Miramar Beach, located in the municipality of Guaymas, Sonora, Mexico, using an integrated approach based on recreational quality and visitor CC.

Technical Summary: The ECC of the beach was determined using the Cifuentes method [16].

The total beach area was derived from satellite images obtained from Google Earth 2016, processed using the software Surfer 10. The ideal values for pedestrian density for beaches used in the equation come from a table suggested by Yepes [70]. The levels Acceptable (5 m2 per person) and Confortable (>10 m2 per person) were selected.

The authors conclude that, even in the busiest season, the ECC is not exceeded, and improving management capability will increase the ECC. No future research work is suggested, but the usage of this method is suggested for similar efforts at other beaches.

Classification:

Table 3.12: Classification of *"Integrated Assessment of Recreational Quality and Carrying Capacity of an Urban Beach"*

Source	DT	PA	SA	TA	V
[47]	B	1	1	P	1

3.2.4.10 Salemi et al.,*"Conceptual framework for evaluation of ecotourism carrying capacity for sustainable development of Karkheh protected area, Iran"*, 2019 [57]

Objective: The objectives of this research were determining the ecotourism potential, identifying the effective indicators in the ecological and social–cultural pressure, assessing the physical, ecological, and social-cultural CC and determining the CC indexes of the Karkheh protected area, Iran.

Technical Summary: The CC of the study area follows the formula of Cifuentes [16].

The total usable area of the ecotourism zones was obtained from the ArcGIS software.

The ecologic CC was calculated by applying ecological correction factors to the PCC, while the social-cultural CC used social and cultural corrective factors. The factors were derived from the analysis of surveys with an Analytic Network Process (ANP) model.

For each CC, a CC index is determined. According to the values of the CC indices, it is concluded the region has an adequate number of visitors for its values of CC.

The authors suggest adopting entry-exit policies for low CC areas, providing instructing sessions and sequential monitoring of the study area.

Classification:

Table 3.13: Classification of *"Conceptual framework for evaluation of ecotourism carrying capacity for sustainable development of Karkheh protected area, Iran"*

Source	**DT**	**PA**	**SA**	**TA**	**V**
[57]	NP	1	1	P	1

3.2.4.11 Blaga et al.,*"Estimating the tourist carrying capacity for protected areas. A case study for Natura 2000 sites from North-Western Romania"*, 2019 [5]

Objective: The main objective of this study was estimating the physical and real tourist carrying capacity for the Natura 2000 Sites: Cefa, Valea Roșie and Ferice Plai.

Technical Summary: The methodology used for the assessment of the PCC of the sites was the Cifuentes [16] method.

The available area for touristic activities was obtained from the management documentations of Natura 2000 sites and aerial photography analysed with multispectral image processing, based on the distribution of the considered tourist resources. The parameters for CC assessment were obtained via surveys. The ECC was not assessed, as there were not quantifiable management capability attributes, according to the authors.

The authors do not suggest future work.

Classification:

Table 3.14: Classification of *"Estimating the tourist carrying capacity for protected areas. A case study for Natura 2000 sites from North-Western Romania"*

Source	**DT**	**PA**	**SA**	**TA**	**V**
[5]	NP	1	2	P	1

3.2.4.12 Moukhtar et al.,*"Assessment of the Environmental Carrying Capacity for Protected Areas: A Study of Petrified Forest and Hassanah Dome, the Great Cairo"*, 2020 [48]

Objective: The study objective is to quantify the environmental CC of Petrified Forest Protectorate in East Greater Cairo and Hassanah Dome Protectorate in West Greater Cairo.

Technical Summary: The methodology used to assess environmental CC is a methodology used by the International Union for Conservation of Nature [15]. The methodology appears to be a reinterpretation of the Cifuentes' methodology [16].

The data and parameters for CC assessment were obtained via surveys and official data.

The indices for the CC of each zone were calculated, using the total population of urban Great Cairo at the time. Both indices indicated an insufficient capacity for tourism activity.

Generally, the research concludes that environmental CC should be taken as a sustainable strategic planning approach to environmental and urban planning processes for any site and any region. As future work, Moukhtar et. al suggest applying the methodology in all protected areas from Egypt.

Classification:

Table 3.15: Classification of *"Assessment of the Environmental Carrying Capacity for Protected Areas: A Study of Petrified Forest and Hassanah Dome, the Great Cairo"*

Source	DT	PA	SA	TA	V
[48]	NP	1	2	P	1

3.2.4.13 Suana et al.,*"Environment carrying capacity and willingness to pay for bird-watching ecotourism in Kerandangan Natural Park, Lombok, Indonesia"*,2020 [62]

Objective: The two objectives of this study are assessing the environmental CC of seven bird-watching ecotourism packages in Kerandangan Natural Park, Lombok, Indonesia, and the willingness to pay (WTP) from visitors for the packages.

Technical Summary: The authors use the Cifuentes method [16] for assessing environmental CC of the following seven ecotourism packages in in Kerandangan Natural Park, Lombok, Indonesia: Main Trail (MT), Southern Hill Trail (SH), Northern Hill Trail (NH), Southern Valley Trail (SV), Northern Valley Trail (NV), Bird Photography (BP) and Night Birding (NB).

The WTP is determined through the collection of data from 150 respondents, obtaining a mean value of US$20.7 per visitor. The obtained values for environmental CC were higher than the current visitors at the time of the study, meaning the bird-watching packages have the means to accommodate more visitors. No future work is suggested.

Classification:

Table 3.16: Classification of *"Environment carrying capacity and willingness to pay for bird-watching ecotourism in Kerandangan Natural Park, Lombok, Indonesia"*

Source	DT	PA	SA	TA	V
[62]	NP	1	2	P	1

3.2.4.14 Makhadmeh et al.,*"Evaluating the carrying capacity at the archaeological site of Jerash (Gerasa) using mathematical GIS modeling"*, 2020 [44]

Objective: The main goal of the present study is to evaluate the CC of the archaeological site of Jerash (Gerasa), in Jordan.

Technical Summary: In this study, the authors use the Boullón methodology [9] and GIS-aided mathematical modeling for assessing a CC map through the whole area of the site, providing a spatially fine-grained estimate according to the monument.

Using the Inverted Distance Weighting, the number of visitors at any unmeasured location was estimated, and with Kernel density, the visitor density map was estimated.

Comparing with an estimated actual visitor density map, the authors analyse the spatial features of the site, concluding some of the zones with the least CC have the most visitors, surpassing their capacities. No future work is suggested, but propose measures for improving the CC management of the site.

Classification:

Table 3.17: Classification of *"Evaluating the carrying capacity at the archaeological site of Jerash (Gerasa) using mathematical GIS modeling"*

Source	DT	PA	SA	TA	V
[44]	AS	1	2	I	1

3.2.4.15 Illiyas et al.,*"Carrying capacity assessment for religious crowd management - An application to Sabarimala mass gathering pilgrimage, India"*,2021 [35]

Objective: This study has the objective of assessing religious crowd CC for the Sabarimala Sree Dharma Sastha Temple pilgrimage, a popular religious site in India, victim of crowd disasters.

Technical Summary: In the present study, the authors apply an adapted version of the Cifuentes methodology [16] for CC assessment. The total usable area is divided into 10 zones, determined via field survey, GIS mapping and space usage observation. The ideal pedestrian density for tourism activity is estimated for each zone. The remaining parameters are obtained via surveys.

It is noted the importance of CC determination and crowd density monitoring for risk-informed crowd management. As future work, the authors suggest the use of the method to develop models and crowd management systems.

Classification:

Table 3.18: Classification of *"Carrying capacity assessment for religious crowd management - An application to Sabarimala mass gathering pilgrimage, India"*

Source	DT	PA	SA	TA	V
[35]	RS	1	2	I,P	1

Table 3.19: Summary table with the related work classification

Source	DT	PA	SA	TA	V
[71]	B	1	1	P	1
[26]	NP	1	2	I	1
[25]	NP;B	1	2	P	1
[19]	B	1	2	P	1
[33]	B	1	1	P	1
[37]	RS	1	2	P	1
[53]	B	1	2	P	2
[56]	NP	1	1	P	1
[47]	B	1	1	P	1
[57]	NP	1	1	P	1
[48]	NP	1	2	P	1
[62]	NP	1	2	P	1
[44]	AS	1	2	I	1
[35]	RS	1	2	I,P	1

3.2.5 Conclusions

The majority of the reviewed studies used the Cifuentes methodology, or a different approach of said methodology, consolidating its validity for assessment of representative CC metrics.

Studies range from a variety of different types of tourism sites, but a big chunk of them focus on the evaluation of beaches. It is clear there are lacking CC studies in urban environments, being an addressable gap in the literature. Naturally, none of the found studies take into consideration CC in a pandemic scenario, as most of the studies were conducted before the COVID-19 outbreak, in late 2019. Usually, when assessing the CC, the spatial grain of the areas is coarse, meaning it spans an entire touristic site. The majority of studies do not account for instant CC, but for the applied areas it may not be as necessary as an urban environment, with a vast variety of different attractions and spatial configuration. Few studies validated their methodologies. Most of them use a single method for CC assessment, lacking the comparison with another method. Nonetheless, the rationale and chosen variables are often validated by surveys, or obtained from another studies.

3.3 Methodology

3.3.1 Data extraction

We start by extracting geographic data from the main OSM webpage[1], using the *export* feature. A bounding box was manually selected to include all data from the Santa Maria Maior district.

Downloading the resulting OSM file, we used the *osmtogeojson* tool [2] to convert the raw OSM XML data into *GeoJSON*[3], a format specification to represent geographic entities in *JSON* format, to insert into a geographic database. Some data transformation was needed, due to an imperfect OSM data representation and transformation to *GeoJSON* with the tool, and some geometries violated the *GeoJSON* specification, needing to be altered.

3.3.2 Data categorization

In order to calculate street areas, we need to utilize these main types of geographic elements:

- **Streets**: Are mainly represented by OSM ways, with an open polyline geometry, with each line representing a street segment. The defining tag is the "Highway" tag, characterizing the way as a street. Its value indicates the type of road.

- **Buildings**: Buildings are represented by nodes, ways or relations. Nodes are not considered, as they cannot be used in polygonal representation, being nodes representative of single points in space. The main defining tag which represents a building is the "building" tag.

- **Open Spaces**: Are urban open places that can be used for pedestrian activity. Examples of the spaces are gardens, grass fields and squares. They are normally represented by polygons, and defined with multiple tags, depending on the space type.

- **Bodies of water**: Bodies of water are places where significant water accumulation of water occurs. Naturally, the area occupied by water is not usable for land tourism activities, needing to be excluded from the total usable areas. These elements are defined by the "water" tag.

3.3.3 Geographic database

The chosen database to store and fetch the required data for the urban zoning determination was the *MongoDB*[4] database. It is a non-relational, document-oriented NoSQL database, which supports *GeoJSON* data handling, native geographic indexing and spatial queries. In order to use these geographic features and increase query performance, it is needed to create indexes on the data collection. The **2dsphere** index provided by the database allows spatial queries on *GeoJSON* documents.

[1] https://www.openstreetmap.org/export
[2] https://github.com/tyrasd/\gls{osm}togeojson
[3] https://geojson.org/
[4] https://www.mongodb.com/

3.3.4 Segment area determination

It is not feasible to manually measure every street and open space in a complex urban fabric, and streets do not have polygonal representations in OSM. To determine the urban fabric zones and calculate the available usable areas, a Python script was developed. Its objective is to generate polygonal representations of street segments through Cartesian operations, contained into the available space of each one, to estimate its area. To be able to perform Cartesian operations, it is needed to convert the latitude and longitude coordinates into Cartesian coordinates. We chose the Universal Transpose Mercator (UTM) coordinate system[5], which divides the globe into sixty east-west zones with 6 degrees of longitude of width, referenced by a number, and each zone is divided by the equator into north and south hemispheres. In each zone, the coordinates are represented by a northing (distance in meters to the north), assigned to the Y-coordinate, and an easting (distance to the east), assigned to the X-coordinate. If an area spans multiple UTM zones or hemispheres, the coordinates in the plane must be adjusted.

For each street, the script iterates its segments. For each segment, a rectangle is created by perpendicularly offsetting the extremities of the segments, until the rectangle intersects a building, square or body of water.

Each segment has two points **P1** and **P2**. To find the perpendicular offsets to both points of each segment, we need to calculate the perpendicular slope to the linear function which represents the segment, defined by

$$y = mx + b$$

where m is the slope of the segment linear function, calculated by the following formula:

$$m = \frac{y2 - y1}{x2 - x1}$$

Being M a perpendicular slope to m, M is defined as the negative inverse of m, being

$$M = -\frac{1}{m}$$

or

$$M = -\frac{x2 - x1}{y2 - y1}$$

The offsetting of the rectangle sides is done by offsetting the segment extremities' coordinates by a step, in the independent coordinate, and by the multiplication of the defined step with the slope, in the dependent variable. This method has the following issues:

- If the slope is too large, a small change in the independent variable results in an excessive change in the dependent variable, creating an outgrown rectangle in a first iteration.
- If the variation of Y is 0, the calculation of the perpendicular slope will result in a *division by zero* error.

To avoid these problems, we evaluate the variations of X and Y, and determine which of the variables has a larger variation. If Y varies less than X ($|y2 - y1| < |x2 - x1|$), M is bigger

[5] https://pubs.usgs.gov/fs/2001/0077/report.pdf

than 1, meaning Y will change more than X, potentially resulting in the previously noted issues. In this case, we change Y to be the independent variable, and X the dependent, inverting the perpendicular linear function to

$$x = \frac{1}{M}y + b$$

The slope of this inverse function is then lower than 1, avoiding the issues. Depending on the selected independent variable, we then define the step of the independent variable as 1 meter, and calculate the step of the dependent variable with the slope of the line. In the first iteration, an initial rectangle is created, adding and subtracting the x steps and y steps to the segment point coordinates. The four vertices A,B,C and D of the rectangle are calculated as follows:

- A = (xP1 + step, yP1 + step Y),
- B = (xP1 - step, yP1 - step Y),
- C = (xP2 - step, yP2 - step Y),
- D = (xP2 + step, yP1 + step Y).

In the next iterations (if there are any), the rectangle will increase in width, further increasing and decreasing the vertex coordinates by the steps, until the sides of the rectangle intersect a surface-level building, a square or a body of water. To do this, the script queries the database for these nearest elements, and then verifies if there is an intersection with any of them. Each side of the rectangle moves independently, to handle streets which are not centered between the surrounding the limiting elements. When one of these elements intersects one of the sides of the rectangle, it stops advancing. A maximum width for the rectangles is defined, in cases where no elements are found nearby, to avoid excessive growth. The resulting polygon areas represent the open area of the urban fabric, and can be used to roughly estimate gross street segment areas. When the rectangles intersect buildings or squares, a percentage of the area which intersects those elements is determined, being considered as unusable area and then subtracted to the total area. If a street segment is entirely contained in a square, it is ignored, as the area of the square is considered as a whole segment. If a segment partially intersects a square, the intersecting square is ignored.

3.3.5 Usable area estimation

With the street segment polygons created, we then need to determine viable methods to estimate the open area that is suitable for pedestrian activity, for each segment. Pedestrian roads will have the totality of their area as usable. Hence, the calculation of the area is estimated as the area of the correspondent segment polygon, subtracting the unusable building area detected in the polygon.

Streets with vehicle roads are not entirely suitable for walking in normal conditions, unless there are events in the area which close the streets for vehicles (e.g., Popular Saints' Festivities in Lisbon). The roads are then blocked during certain periods of time or special permits are required to be traversed by vehicles (e.g. for loading and unloading goods), as it occurs in some streets in Santa Maria Maior neighborhood.

If the streets are not classified as pedestrian streets, the usable area is limited to the sidewalk area. OSM data coverage of sidewalk width is poor (non-existent in the district), so we need to estimate sidewalk area by subtracting the total segment's area by the estimated unusable area. To estimate the unusable area in non-pedestrian roads, we calculate the estimated vehicle road area. This area is estimated by multiplying the segment height by the total vehicle road width, which is estimated by multiplying the number of lanes by reference lane width values used for each road type. The chosen value for default lane width is the minimum road width allowed in Lisbon's municipality's master development plan (PDM) [6] and Lisbon's Public Space Manual [12], which establishes the minimum lane width for a road as 3 meters.

Street parking space is also considered unusable area, and is estimated by analysing the number and type of parking lanes in the street's data. If a parking lane is defined as diagonal or perpendicular, we use the average value of 5 meters for the width of parking lane, and 2 meters for the width of parallel parking lanes. The estimated parking lane width is then multiplied by the length of the street segment, obtaining the parking area of the segment. In case of missing relevant information, assumptions are needed for a more accurate estimation. Information about street parking is missing in many streets in Santa Maria Maior. Observing multiple streets from the district, we define an assumption that a street with a single lane and more than 6m of average width to have an additional single parallel parking lane.

3.3.6 PCC assessment

Having the urban tourism usage zoning and usable area determined, the PCC of each zone can be estimated. The used method for assessing PCC of each zone is an adapted version of the Cifuentes methodology [16], removing the temporal factor (rotation factor), and converting it to the PCC of the zones of an instant of time. The equation for assessing the instant PCC of a zone is the following:

$$PCC = \frac{A}{Ap}$$

where A is the usable area of the zone, and Ap is the ideal free available area by pedestrian, for comfortable tourist activity of the urban places. We will calculate the PCC for each highway pedestrian usage level of service (LOS) for moving pedestrians, defined in the Highway Capacity Manual [54], and adapted in Lisbon's Public Space Manual [12]. Each LOS is defined by a range of available space per pedestrian (Ap), pedestrian flow and average walking speed, and allows to categorize different levels of crowding in the public urban space. Each LOS also has a description of the general confort for the pedestrians. For the assessment of the PCC thresholds, only the available space per pedestrian of each LOS will be considered.

In Table 3.20, the characteristics of each LOS are presented.

[6] https://www.lisboa.pt/fileadmin/download_center/normativas/regulamentos/urbanismo/Regulamento_PDM.pdf

Table 3.20: Pedestrian LOS description (adapted from [12])

LOS	Space per pedestrian (m2/person)	Comfort
A	> 5.6	Very high comfort
B	3.7 – 5.6	High comfort
C	2.2-3.7	Medium comfort
D	1.4-2.2	Low comfort
E	0.75-1.4	Very low comfort

To account for pandemic conditions (COVID-19 pandemic), a minimum social distancing of 2 m between persons was advised by the *Direção Geral de Saúde* (DGS)[7]. The LOS B is the maximum LOS which allows, on average, for social distancing to occur across the public space. For the assessment of pandemic-aware PCC, we define *Ap* as 4 m2 per pedestrian, as it allows for the recommended 2 m of interpersonal distance. For normal conditions, We will consider the as *Ap* as the lower limit of LOS C (2.2 m2 per pedestrian), being the resulting PCC the value of threshold of overcrowding, based on the level of comfort it represents.

To calculate the PCC of areas composed with multiple segments (e.g., streets), the capacities of each segment belonging to that area are summed, trimming possible outlier polygons, which can occur, e.g., in street intersections. For the assessment of the CC of squares, the script sums the capacity of the original OSM polygonal representation of the spaces with all segments of the surrounding streets. The squares of the district are inconsistently represented, as some squares have a mapping of a polygon with the space limited by the surrounding buildings, and others only have the central pavement, which is surrounded by roads, and do not represent the entire area of the square.

3.4 Results and validation

The script is able to successfully divide the open space of Santa Maria Maior into polygonal zones. In Figure 3.2, we present a map, showing the buildings and the resulting zone segments of the district. We can observe an adequate coverage of the usable open space of the complex urban fabric.

For an initial validation of the results, we select 10 streets in the district from the *Pombaline Downtown*, which have different and comparable characteristics (e.g., average width, length, road type, width irregularity) and cover a decent amount of street scenarios. We also selected 4 well-defined squares for validation of open space area estimation. For each street and square, we compare the total area (without building area) calculated by the script with manually measured values using *geojson.io*[8], creating *GeoJSON* polygonal representations of the selected streets and calculating their areas, and calculate the percent accuracy. The average percent accuracy for the streets and for the squares is also assessed.

The validation of the usable area estimation is not in the scope of this dissertation, as it requires in-field measurement for accurate accepted values.

[7] https://covid19.min-saude.pt/wp-content/uploads/2020/04/Distanciamento-social-07-04-2020.pdf

[8] https://geojson.io/

Figure 3.2: Resulting zoning polygons (in red) in the Santa Maria Maior

Tables 3.21 and 3.22 show the estimated total areas, total unusable areas, total usable areas, normal PCC and COVID-19 PCC for the selected streets and squares, respectively.

In Tables 3.23 and 3.24, the estimated total area, the *geojson.io* measured total area and percent error for the selected streets and squares are displayed, respectively.

Table 3.21: Streets with the calculated values for CC

Street	Total Area (m2)	Unusable Area (m2)	Total Usable Area (m2)	Normal CC (persons)	COVID-19 CC (persons)
Calçada de São Francisco	1874.57	1056.45	818.12	371	204
Rua Augusta	7272.03	0	7272.03	3305	1818
Rua da Assunção	2368.5	175.42	2193.08	996	548
Rua da Conceição	3165.76	1766.61	1399.15	635	349
Rua da Prata	7971.18	3418.99	4552.19	2069	1138
Rua da Vitória	3001.77	0	3001.77	1364	750
Rua de São Nicolau	3284.1	180.31	3103.79	1410	775
Rua dos Correeiros	3092.76	0	3092.76	1405	773
Rua dos Sapateiros	3002.9	1946.18	1056.72	480	264
Rua Áurea	8209.25	3550.84	4658.41	2117	1164

Table 3.22: Squares with the calculated values for CC

Square	Total Area (m2)	Unusable Area (m2)	Total Usable Area (m2)	Normal CC (persons)	COVID-19 CC (persons)
Praça da Figueira	10382.57	1749.76	8632.81	3904	2140
Praça Dom Pedro IV	16912.14	3500.13	13412.01	6077	3331
Praça do Comércio	36606.12	778.53	35827.58	16281	8952
Praça Martim Moniz	20910.11	3631.43	17278.68	7834	4294

Table 3.23: Streets with the measured total areas, estimated total areas and accuracy

Street	Measured Total Area (m2)	Estimated Total Area (m2)	Accuracy (%)
Calçada de São Francisco	1640.71	1874.57	85.7
Rua Augusta	7262.96	7272.03	99.9
Rua da Assunção	2136.97	2368.5	89.2
Rua da Conceição	2918.39	3165.76	91.5
Rua da Prata	7400.98	7971.18	92.3
Rua da Vitória	2834.98	3001.77	94.1
Rua de São Nicolau	2781.19	3284.1	81.9
Rua dos Correeiros	3072.15	3092.76	99.3
Rua dos Sapateiros	2939.56	3002.9	97.8
Rua Áurea	7059.11	8209.25	83.7
Average accuracy			**91.5**

Table 3.24: Squares with the measured total areas, estimated total areas and accuracy

Square Name	Measured Total Area (m2)	Estimated Total Area (m2)	Accuracy (%)
Praça Dom Pedro IV	18897.84	16912.14	89.5
Praça Martim Moniz	24660.1	20910.11	84.8
Praça da Figueira	12016.27	10382.57	86.4
Praça do Comércio	35596.45	36606.12	97.2
Average accuracy			**89.5**

Analysing the resulting data, we can observe an average of area estimation accuracy of 91.5% for the selected streets, and 89.5% for the selected squares. Higher error percentages in some streets can be explained by street limiting differences between the polyline representation of the street and the polygonal representation (frontiers with buildings). Exact representations of the frontiers of streets and squares can be hard to define. Since PCC is used as an initial reference value to measure overcrowding, and due to its imperfect nature, the obtained error is acceptable for the assessment of PCC. In chapter 5, we assess the feasibility of the usage of the calculated segments and CC thresholds to evaluate the crowding state in the Alfama neighborhood.

3.4.1 Validation threats

The accuracy of the algorithm heavily relies on the accuracy of the mapping in the OSM project. Since this mapping project is backed by volunteers, which can be anyone, some data inaccuracies can occur, and relevant data can be missing. Having a lack of relevant data for usable area estimation requires assumption-based variable estimations, which can lead to less accurate results.

The manual area measurement using *geojson.io* is prone to human error, and the precision of the measurement tool is unknown, which can lead to accuracy errors.

The map used in *geojson.io* for the manual polygon creation is based on OSM data, which is the base data for the script. Ideally, other source of geographic data could be used, to validate against OSM data errors.

The number of selected streets and squares that were used to validate the results may be less than ideal. The sample may not represent the average accuracy of the algorithm on all the streets and squares in the district.

Chapter 4.

Public Space Mass Event Egress

Contents

This chapter presents the motivation, related work, the agent-based model to simulate public space non-urgent egress, and results

[This page has been intentionally left blank]

4.1 Motivation

Cities have integrated in their tourism development strategies the realization of crowded events in public spaces, such as music festivals, fireworks and video mapping. Other public events such as political rallies, protests, celebrations or commemorations can attract an immense number of visitors. In Figure 4.1 is shown a crowded event in a public open space.

Figure 4.1: New Year's Eve celebration at Praça do Comércio square, Lisbon[1]

Choosing the adequate public space for each of these events requires knowing the expected number of visitors, acoustic conditions, spatial context and accessibility. For the success of these events, adequate planning is required for the allocation of security, emergency and cleaning resources. Thus, it is necessary to undertake risk analysis based on the carrying capacity and egress time from the public space hosting the event. The safety of urban public space events depends on a wide set of factors, such as the scope of the event, the characterization of the expected visitors and the geometry of the space. For instance, a protest of senior citizens would have different risk factors as a concert for younger crowds. A rally for a tolerance preaching political party would not be as risky as an extremist radical party rally. A quick event would probably not potentiate as much damage as a long duration event. In case of the appearance of order disrupting situations, such as natural disasters, explosions or terrorist attacks, the geometry of the public space must allow for quick and efficient evacuation, in order to avoid or keep casualties to a minumum.

To study the impact of spatial configuration of open public spaces for the egress of pedestrians, we will consider three widely used public spaces for crowd event hosting, with similar areas, in Lisbon: *Praça do Comércio*, *Praça Marquês de Pombal* e the east side of *Alameda Dom Afonso Henriques*. Through the use of OSM data and the GAMA platform, we developed an agent-based model to simulate the egress dynamics of pedestrians.

[1] https://www.culturanarua.pt/noticia/ano-novo-e-no-terreiro-do-paco/

4.2 Related work

In this section, we will present a literature review to analyse relevant state-of-the-art work done on public space egress simulation, following the previously explained Rapid Review methodology.

4.2.1 Research objectives

The objectives of this literature review are:

- Undertake a rapid systematic review of research on public crowd space egress modeling and simulation;
- Filter the articles by relevancy to the given topics to further review in more detail;
- Integrate information from these studies;
- Determine how the various topics approached in the reviews relate to each other through taxonomic classification;
- Identify opportunities for improvement/further investigation.

Concretely, we want to determine:

1. What is the current body of knowledge about the approaches on public crowd space egress modeling and simulation?
2. What are currently the most effective approaches used for public crowd space egress modeling and simulation?

4.2.2 Review methodology

4.2.2.1 Study selection criteria

Selected articles must present empirical data or the production of publicly available, usable and testable artifacts. The inclusion of studies will not be restricted to any specific type of intervention. Real case or hypothetical scenarios are included.

4.2.2.2 Inclusion and exclusion criteria

To include or exclude found articles in this review, the defined criteria are in Tables 4.1 and 4.2

Table 4.1: Inclusion criteria

Inclusion criterion	Description
IC1	Studies must be published in journals after 2011
IC2	Articles must be related to the assessment of PCC of touristic places of interest
IC3	Articles must be research-based
IC4	Studies must be peer-reviewed

Table 4.2: Exclusion criteria

Exclusion criterion	Description
EC1	Articles are inaccessible
EC2	Articles are not written in English

4.2.2.3 Search strategy

The adopted strategy was to search related work on the *Scopus* scientific database, performing the forward and backward snowballing procedure [61].

4.2.2.4 Search strings

Search strings were created through trial and error, combining multiple terms related to the domains of each study, as explained in the previous literature review. The *TITLE-ABS-KEY* operator indicates that the search engine only checks for the search strings in the title, abstract and keywords of the studies. The *PUBYEAR* term filters for documents published before or after a given year, depending on the comparison symbol used (< or >). The search strings used to search on the database are the following:

- TITLE-ABS-KEY ((pedestrian OR visitor OR person) AND egress AND (event OR exit) AND simulation AND NOT evacuation) AND PUBYEAR > 2010

4.2.2.5 Selection process

The selection process steps are defined as:

1. Execute the search string on the defined databases and search engines;
2. Apply the inclusion criteria (orderly by title, abstract, keywords and content);
3. Apply the exclusion criteria;
4. Apply backward and forward snowballing to the remaining articles;
5. Repeat steps 2 and 3 on the new articles selected by snowballing.

4.2.2.6 Validity threats

This review has possible threats to its validity. First, the number of articles used for the analysis of the related work may be less than ideal to assure the validity of the results. Out of all the related work found, only a total of 4 articles were selected.

First, only using a single search engine may affect validity, as there may be useful related studies which have not been properly indexed, missing on them.

Second, more variations of the search strings, such as appending more terms and varying term combinations, could have been used to narrow the database results to a lesser number of more relevant articles. Limiting the search strings might also have led to not finding some related work.

Last, inaccessible articles were discarded, which could lead to the loss of important information. An effort to reach the authors should have been done.

4.2.3 Classification taxonomy

4.2.3.1 Taxonomy criteria

To categorize the different types of pedestrian egress simulation, a set of criteria was defined. Table 4.3 shows the selected criteria.

Table 4.3: Taxonomy criteria

Criterion	Acronym
Space continuity	SD
Scenario realism	SR
Modelling scale	MS
Egress urgency	EU
Space type	ST
Validation	V

4.2.3.2 Space Continuity (SC)

Egress models can have environments modeled in discrete space or continuous space. The criterion categorizes the studies by the following nominal scale:

- **Discrete Space**(DS): The space of the environment is discrete;
- **Continuous Space**(CS): The space of the environment is continuous.

4.2.3.3 Scenario Realism (SR)

The selected studies can have egress simulations to describe real-life places through static digital twinning, or simply study pedestrian dynamics in fabricated scenarios. This criterion categorizes the simulation scenario realism by the following ordinal scale:

- **Abstract Environment**(AE): The modeled environment is fabricated, and does not represent a real-life environment;
- **Real Environment**(RE): The modeled environment in simulations represents a real-life location.

4.2.3.4 Modelling Scale (MS)

Modelling can be divided into two scales: microscopic and microscopic. A microscopic model considers the interaction of individual agents, while a macroscopic model considers a higher level collective behavior of multiple entities. The criterion categorizes the models in the studies by the following nominal scale:

- **Microscopic (MIC)**: The model is microscopic;

- **Macroscopic (MAC)**: The model is microscopic.

4.2.3.5 Egress Urgency (EU)

The egress models may describe scenarios of threat-induced, emergency evacuation, or non-urgent normal egress. The criterion categorizes by a nominal scale as follows:

- **Urgent (U)**: The model represents urgent scenarios;
- **Non-Urgent (NU)**: The model represents non-urgent scenarios.

4.2.3.6 Space Type (ST)

Egress models can describe indoor or outdoor scenarios. This criterion categorizes by the following nominal scale:

- **Open space (OS)**: The modeled environment represents an open, outdoor space, such as a city square.
- **Enclosed space (ES)**: The modeled environment in simulations represents enclosed environments, such as rooms in buildings, train stations, etc.

4.2.3.7 Validation (V)

This criterion describes if a study is validated or not. The following ordinal scale categorizes the criterion:

- **Not Validated (NV)**: The study is not validated.
- **Validated (V)**: The study is validated.

4.2.4 Related work review

4.2.4.1 Heliövaara et al.,*"Patient and impatient pedestrians in a spatial game for egress congestion"*,2013 [31]

Objective: The main objective of the present study is to present an agent-based spatial game to model the interaction of agents in situations of egress congestion due to the interaction of patient and impatient pedestrians.

Technical Summary: The authors derived a microscopic agent-based spatial game model, where each agent develops strategies in order to maximize their reward, playing with another agents.

The agents act on a cost function that describes the risk of not evacuating in time, which causes the agents to play a different game depending on their distance to the exit. Each agent interacts with the neighbouring agents, and can have 2 behavioral strategies: impatient and patient behavior. The impatient agents try to push forward and overtake other agents, while patient agents move with the crowd and want to avoid physical contact. The choice of the behavior will depend on the result of the played game between agents. The equilibria of the game are calculated with different parameter values in continuous and discrete spatial settings. The agents update their strategies frequently based on their best-response functions, which only react to the current strategies of their neighbors, and optimize the behavior given their condition. The game model is implemented with the classic SFM [30] to model agent movement. The agents select their strategies from the game model by making the individual movement parameters depending on the strategy. The model was simulated in scenarios of fixed and adaptive strategies. The results have shown that threat-induced clogging may be caused by the rationally of crowd members, instead of the idea that it is caused by panic irrationality. No future research work is suggested.

Classification:

Table 4.4: Classification of *"Patient and impatient pedestrians in a spatial game for egress congestion"*

Source	SC	SR	MS	EU	ST	V
[31]	DS;CS	AE	MIC	U	ES	1

4.2.4.2 Wagoum et al.,*"Understanding human queuing behaviour at exits: an empirical study"*,2017 [68]

Objective: The goal of this study is to perform an empirical study on the egress exit choice behavior of pedestrians, using laboratory pedestrian experiments and model simulation.

Technical Summary: In this study, 3 groups of laboratory experiments were performed with the BaSiGo framework [32], with real pedestrians in different symmetric and asymmetric geometry configurations. The experiment results have shown the occurrence of load balancing of the distribution of the participants over the exits, resulting in optimized egress times, and in normal scenarios, the exit choice in not only based on the distance to the exits.

From the performed experiments, 2 exit choice models were developed: One where the exit choice is a function to the distance of the exits, and other with a function of distance to the exits and the density of pedestrians in front of the exits. The parameters for the models were estimated and calibrated through a linear discriminant analysis. The trajectories were modeled and simulated with microscopic pedestrian models presented in [65]. Simulations with the models were performed in 3 experiments, analogous to the real experiments.

The results suggest the distance-based model describes unrealistic behaviours in the case of asymmetric scenarios or congestion, and load balancing does not occur. The distance-density model can describe the load balancing of the pedestrians over the exits. Visibility constraints and nonlinear effects can affect model precision. Knowledge of the environment, preferences on herding, inertial system and relaxation are mechanisms that can affect model accuracy, and require additional parameters not considered in the present models. As future research work, the authors propose the study of scenarios with larger areas and crowd densities.

Classification:

Table 4.5: Classification of *"Understanding human queuing behaviour at exits: an empirical study"*

Source	SC	SR	MS	EU	ST	V
[68]	CS	AE	MIC	NU	ES	1

4.2.4.3 Vizzari et al.,*"An agent-based model for plausible wayfinding in pedestrian simulation"*,2020 [67]

Objective: The purpose of the present study is to present a model to integrate consolidated results on pedestrian behavior modelling at the basic locomotion level with higher level path decisions in complex, multi-passage environments.

Technical Summary: This work improves the previous work by the authors, on the definition of a model for wayfinding decisions simulation considering perceivable congestion in path planning [17], and on the introduction of an imitation mechanism, and calibration of the significance of the path length, perceived congestion and perceivable recent changes in the model [18], providing a deeper assessment of commonsense evaluation of the effects of perceivable congestion on path planning.

The presented model is a microscopic ABM in discrete environment space, modeled with a square cell grid with 40cm sided cells, considering the average space occupied by a pedestrian. It uses an extended version of the floor fields approach by [11] for agent navigation.

The decision-making of pedestrian agents is based on the following observations: (i) pedestrians can perceive congestion when planning their paths; (ii) have imprecise reasoning due to limited decision time and imprecise individual perception and (iii) are influenced by other surrounding pedestrians through imitation mechanisms. This allows agents to consider distances and the evaluation of the current simulation dynamics.

The validation and calibration of the model is done by conducting experiments with real pedestrians in a controlled setting, and sensitivity analysis.

The model was simulated in a realistic scenario of the egress from a football arena, allowing the analysis of space utilization and path plan reconsideration.

The results show that the present model is plausible but not optimal. The achieved insights from the study could influence researches on other composite decision-making activities. Although the model was implemented in a discrete space approach, the authors point it can be implemented in continuous space approaches.

As future work, the authors indicate the incorporation pedestrian group dynamics in wayfinding decisions, developing a serious game, and employ video game and virtual reality technologies to acquire sustainable empirical evidence to further validate the models.

Classification:

Table 4.6: Classification of *"An agent-based model for plausible wayfinding in pedestrian simulation"*

Source	SC	SR	MS	EU	ST	V
[67]	DS	AE	MIC	NU	ES	1

4.2.4.4 Zhuang et al.,*"Exploring the behavior of self-organized queuing for pedestrian flow through a non-service bottleneck"*, 2021 [73]

Objective: The present study presents an agent-based cellular automata model to describe the organized queuing at non-service bottlenecks, where agents are able perceive and act from the social environment order in real time.

Technical Summary: Zhuang et.al. adapted social-psychological findings to a agent-based cellular automata model that describes self-organized queuing before bottlenecks in non-panic situations. A perceive–act mechanism is presented, serving as base for the agents' behavior. The social system is based on the completion of 4 functional problems by each agent: (1) goal attainment, (2) adaptation, (3) integration and (4) latency.

A set of simulations in a small, abstract environment were conducted, which are compared with real-life controlled queuing experiments. The flow–density, flow–velocity relationships and time-varied perception of the group order can be successfully simulated.

The results of the simulation shown an extremely high-ordered environment is not favorable for collective egress, as the entering process unfairness and local congestion at the queue end greatly reduce the average speeds of pedestrians, thus, a moderate orderly environment can be more effective for alleviating local congestion.

As future work, the authors propose virtual reality laboratory experiments to further calibrate and validate the proposed methodology, including physiological arousal to verify the subjective ratings of observers, the use of heterogeneous subjects, to extend to other queuing cultures, and providing suggestions for more general policy-making and crowd management to improve order regulation awareness for the benefit of bottleneck passing.

Classification:

Table 4.7: Classification of *"Exploring the behavior of self-organized queuing for pedestrian flow through a non-service bottleneck"*

Source	SC	SR	MS	EU	ST	V
[73]	DS	AE	MIC	NU	ES	1

Table 4.8: Summary table with the related work classification

Source	SC	SR	MS	EU	ST	V
[31]	DS;CS	AE	MIC	U	ES	1
[68]	CS	AE	MIC	NU	ES	1
[67]	DS	AE	MIC	NU	ES	1
[73]	DS	AE	MIC	NU	ES	1

4.2.5 Conclusions

Although there is relevant work about the modelling and simulation of pedestrian egress, there is a heavier focus in urgent evacuations, and a smaller emphasis in non-panic egress.

Most of the found studies address fabricated egress scenarios in buildings and other indoor structures, with a relatively small number of agents. Thus, there is a gap in the literature to address massive amounts of agents in mass events in large open spaces. Both cellular automata models, with discrete space environments and continuous space models are also used successfully in these studies, as well as microscopic models.

Concrete egress simulations on environment digital twins of real-life locations were not found. All the reviewed studies had fabricated environments.

The classic SFM is very relevant in egress studies, and pedestrian studies in general. The model is still applied in many models, with slight or heavy modifications.

4.3 Agent-based model

This simple ABM is developed to model and simulate the non-urgent egress of visitors from mass gathering events in public open spaces, namely, squares or plazas, in order to understand how the different built environments of the open spaces, crowd density and exit choosing scenarios affect the egress times of the event visitors. The environment is modeled after the open urban open space using OSM building geometry data and manually designed geometries. The leaving pedestrians are modeled as autonomous agents who interact with each other and the environment in order to progress towards their desired exits, simulating crowd maneuvering. The simulations of the model in various scenarios will be run on a remote Virtual Machine (VM), generating data files and simulation snapshots for the analysis of the results.

This work is a continuation of our previous work [10], where we analysed the egress times for multiple scenarios of crowd density in the open spaces using a derived model for egress flow rates in exits, from the *Society of Fire Protection Engineers (SFPE) Handbook* [34].

4.3.1 Simulation software

As previously stated in the dissertation, the used ABM platform is the GAMA 1.8.2 alpha build.

4.3.2 Environment modelling

The model environment is composed by the buildings that surround and delimit the open space, as well as the buildings inside (e.g., statues, monuments), the usable area of each public open space, which is modeled as a polygon enclosed by the surrounding buildings or borders with bodies of water (e.g., river boundaries, coast), where pedestrians will randomly spawn, and the exit geometries, which are modeled as polygons that cover the entrances to the exiting streets, bordering the open space polygons. We used QGIS 3.16 LTS to manually model the exit geometries and open space limits polygon, and to download and process OSM building geometry data to model the built environment of the public open spaces. Figures 4.2, 4.3 and 4.4 show aerial pictures of the modeled squares, obtained from Google Maps[2], alongside snapshots of their model counterparts from GAMA. The dark gray geometries represent the limiting and internal buildings and other open space delimiters, and the red geometries represent the modeled exits from the spaces.

[2] https://www.google.pt/maps

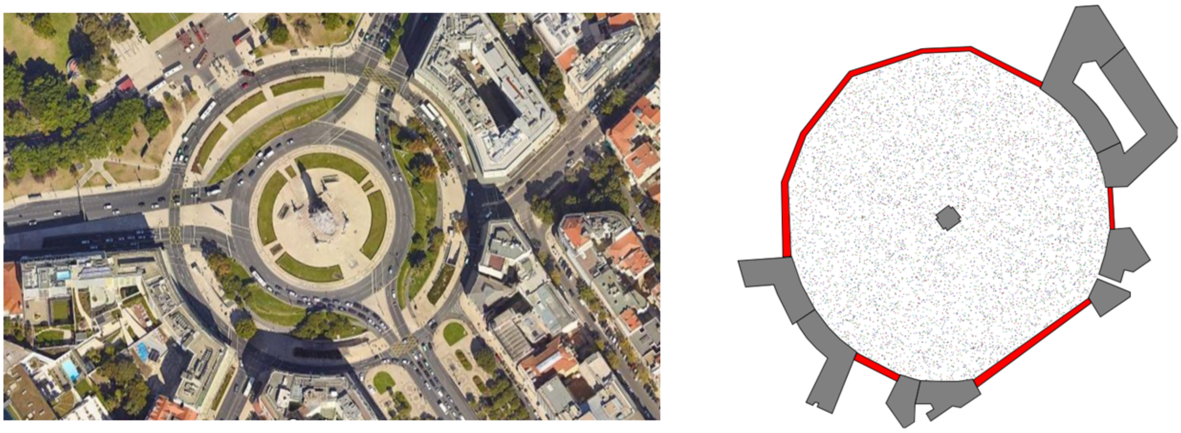

Figure 4.2: Aerial picture of the Marquês de Pombal square (left) (source: Google Maps), and a snapshot of the modeled square in GAMA (right).

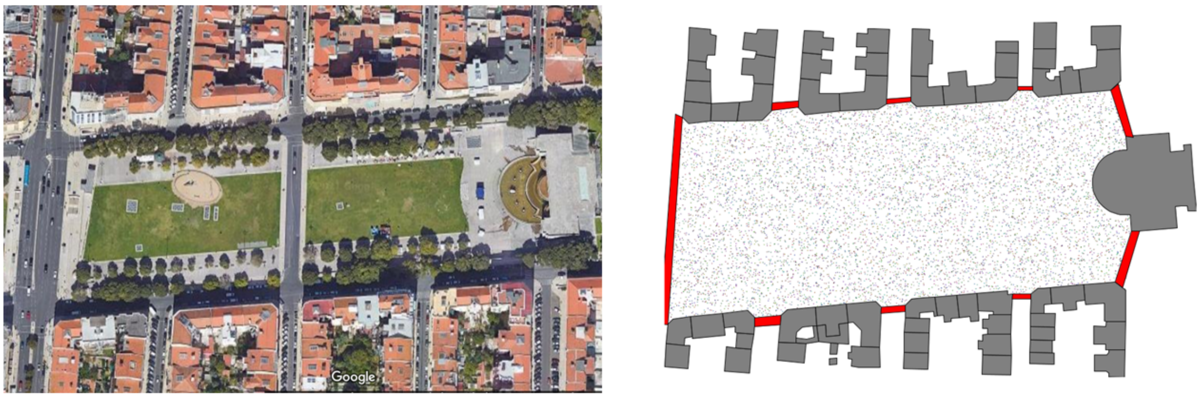

Figure 4.3: Aerial picture of the eastern side of the Alameda square (left) (source: Google Maps), and a snapshot of the modeled square in GAMA (right).

4.3.3 Agent modelling

The only active agents in the model are the pedestrians who egress from the open spaces. These agents walk towards the chosen exit using the built-in Pedestrian skill[3], based on Helbing and Molnar's SFM [30], evaluating the repulsive social forces other pedestrians and obstacles act upon them. The building geometries are also considered agents, but do not have behavior, serving as obstacles for the SFM calculations by the Pedestrian skill.

4.3.3.1 Agent behavior

In the beginning of the simulation, the number of defined pedestrians is spawned in uniform distribution in random locations in the defined open space polygons. When the pedestrians are created, they choose an available exit. Depending on the simulation scenario, the pedestrian chooses the nearest exit, a random exit with uniform random probability and random probability weighted by exit width.

At each cycle, a pedestrian agent behaves as follows:

[3] https://github.com/gama-platform/gama/blob/GAMA_1.8.2/miat.gaml.extensions.pedestrian

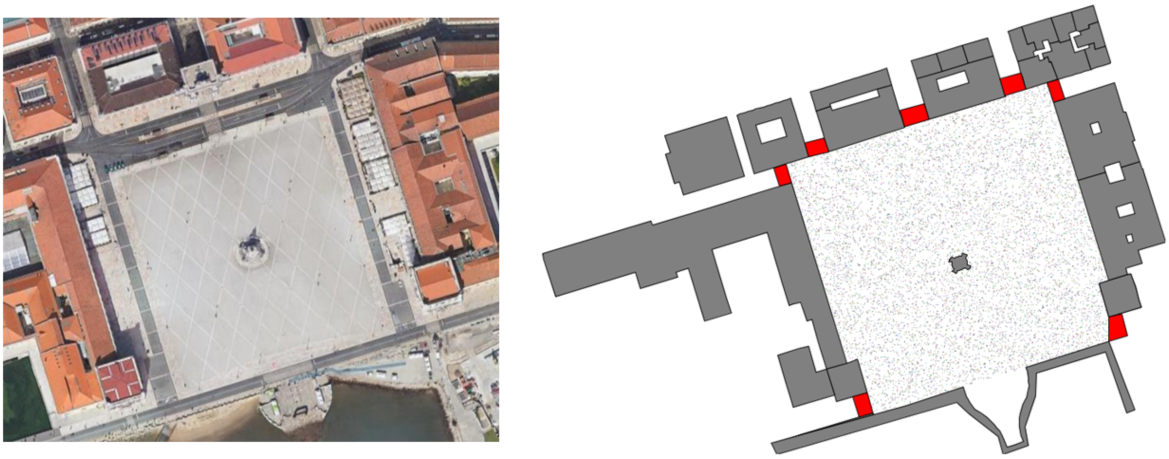

Figure 4.4: Aerial picture of the Praça do Comércio square (left) (source: Google Maps), and a snapshot of the modeled square in GAMA (right).

1. The agent calculates the sum of the social forces acting upon it. If obstacles and/or pedestrians are in the consideration distance to the agent, they will exert social forces on the agent. The final speed and direction of the agent are calculated from the resulting force;

2. The agent moves according to the calculated speed and direction;

3. If the agent reaches its chosen exit, it is removed from the simulation. The variable containing the number of remaining pedestrians is decremented.

In Figure 4.5 is shown a close-up snapshot of a running simulation. Each triangle represents a pedestrian agent, moving towards the closest exit, represented by the red geometry.

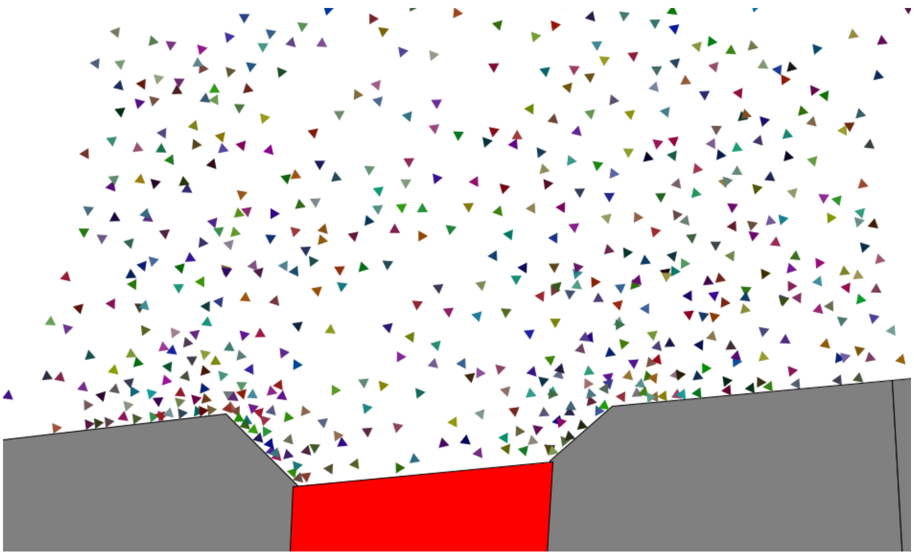

Figure 4.5: Close-up snapshot of pedestrian agents reaching one of the exits of Alameda

4.3.3.2 Pedestrian skill parameters

The values for the parameters for the Pedestrian skill are the default *"classic"* values in GAMA. In Table 4.9 are displayed the set values for the variables.

Table 4.9: Pedestrian skill SFM parameters, the description of the parameters from the GAMA documentation and value

SFM Parameter	Description from GAMA documentation	Value
shoulder_length	The width of the pedestrian (in meters) - classic values: [0.39, 0.515]	0.45
A_pedestrians_SFM	Value of A in the SFM model for pedestrians - the force of repulsive interactions (classic values : mean = 4.5, std = 0.3)	4.5
A_obstacles_SFM	Value of A in the SFM model for obstacles - the force of repulsive interactions (classic values : mean = 4.5, std = 0.3)	4.5
B_pedestrians_SFM	Value of B in the SFM model for pedestrians - the range (in meters) of repulsive interactions	2.0
B_obstacles_SFM	Value of B in the SFM model for obstacles - the range (in meters) of repulsive interactions	2.0
k_SFM	Value of k in the SFM model: force counteracting body compression	200
kappa_SFM	Value of kappa in the SFM model: friction counteracting body compression	400
relaxion_SFM	Value of relaxion in the SFM model - the amount of delay time for an agent to adapt .(classic values : mean = 0.54, std = 0.05)	0.54
gama_SFM	Value of gama in the SFM model the amount of normal social force added in tangential direction. between 0.0 and 1.0 (classic values : mean = 0.35, std = 0.01)	0.35
lambda_SFM	Value of lambda in the SFM model - the (an-)isotropy (between 0.0 and 1.0)	0.5
n_SFM	Value of n in the SFM model (classic values : mean = 2.0, std = 0.1)	2.0
n_prime_SFM	Value of nín the SFM model (classic values : mean = 3.0, std = 0.7)	3.0

4.3.4 General assumptions

The general assumptions for the egress models are the following:

- Pedestrians are created with walking speeds following normal distribution with an average of 1.34 m/s and a standard deviation of 0.26 m/s [30];
- Group behavior is neglected. Every agent is considered alone, and doesn't follow other agents;
- Every pedestrian has the intention to leave immediately. It may not be realistic, as in some cases visitors may not be forced to leave the event space after its end, and then may choose to stay in the space;
- All area without buildings inside the limit polygon is considered usable area. In urban open space mass events, road are often blocked to vehicles, only allowing pedestrians or special vehicles traffic (e.g., ambulances, police cars, event staff vehicles);
- Pedestrian flow beyond the exits is not considered, being assumed as non restricting to egress flow. The blockage ahead of the exits is ignored, which may affect egress rates.

4.4 Simulation scenarios

The model will be simulated in a total of 18 scenarios. For each of the 3 studied Lisbon open spaces, we will simulate the egress with event capacities of 3 pedestrians per m2 (maximum conditions) and 0.25 pedestrians per m2 (COVID-19 conditions). For each of these scenarios, we will simulate 3 different exit choice behaviors: choosing the closest exit (***Closest***), choosing a random exit with uniform probability between each exit (***Random***) and choosing a random exit with weighted probability by exit width (***Weighted Random***). These exit choosing behaviors are not descriptive of real-life scenarios, but represent plausible sub-optimal egress conditions, allowing to evaluate the handling of worst-case scenarios of the public open spaces. The defined step for each simulation cycle is 1 second. The simulations end when the percentage of remaining pedestrians in the open space reaches 10% (90% egress time). Each scenario will be run 10 times to evaluate the influence the randomization of variables can generate. As output of each simulation, a CSV (Comma Separated Values) file with the remaining pedestrians in each passed second in model time, and snapshots of the simulation visualization for each simulation cycle.

In Table 4.10, the pedestrian capacities for each open space, event condition and estimated total exit width is shown.

Table 4.10: Egress scenario conditions by open space

Public space	**Area (m2)**	**Maximum capacity (persons)**	**COVID-19 capacity**	**Total exit width (m)**
Praça Marquês de Pombal	31261	93784	7815	392
Alameda Dom Afonso Henriques (east)	33824	101473	8456	254
Praça do Comércio	35596	106789	8899	108

4.5 Results

In Table 4.11, the results for the run simulations can be observed.

Table 4.11: Average and standard deviation of the egress time (in seconds) for each open space egress scenario

<table>
<tr><th>Square</th><th>Crowd Density</th><th>Exit choice behavior</th><th>Average 90% Egress Time (s)</th><th>Standard Deviation</th></tr>
<tr><td rowspan="6">Praça Marquês de Pombal</td><td rowspan="3">Maximum capacity (3 person / m2)</td><td>Closest</td><td>179</td><td>0.00</td></tr>
<tr><td>Random</td><td>643</td><td>1.41</td></tr>
<tr><td>Weighted Random</td><td>417</td><td>0.00</td></tr>
<tr><td rowspan="3">COVID-19 capacity (0.25 person / m2)</td><td>Closest</td><td>88</td><td>0.00</td></tr>
<tr><td>Random</td><td>295</td><td>2.24</td></tr>
<tr><td>Weighted Random</td><td>197</td><td>1.00</td></tr>
<tr><td rowspan="6">Alameda Dom Afonso Henriques (east)</td><td rowspan="3">Maximum capacity (3 person / m2)</td><td>Closest</td><td>185</td><td>0.00</td></tr>
<tr><td>Random</td><td>870</td><td>2.83</td></tr>
<tr><td>Weighted Random</td><td>554</td><td>0.00</td></tr>
<tr><td rowspan="3">COVID-19 capacity (0.25 person / m2)</td><td>Closest</td><td>105</td><td>0.00</td></tr>
<tr><td>Random</td><td>397</td><td>3.61</td></tr>
<tr><td>Weighted Random</td><td>276</td><td>2.45</td></tr>
<tr><td rowspan="6">Praça do Comércio</td><td rowspan="3">Maximum capacity (3 person / m2)</td><td>Closest</td><td>287</td><td>0.00</td></tr>
<tr><td>Random</td><td>788</td><td>1.73</td></tr>
<tr><td>Weighted Random</td><td>535</td><td>1.00</td></tr>
<tr><td rowspan="3">COVID-19 capacity (0.25 person / m2)</td><td>Closest</td><td>162</td><td>1.41</td></tr>
<tr><td>Random</td><td>373</td><td>2.65</td></tr>
<tr><td>Weighted Random</td><td>264</td><td>2.00</td></tr>
</table>

For the selected squares, the areas and similar, meaning that for each crowd density scenario across all squares, the number of pedestrians will be close to each other, making them comparable in terms of spatial topology.

For a total of 10 simulations for each scenario, the standard deviation of the samples in most scenarios is close to 0, with a few scenarios reaching up to 3.61. This suggests the emergent phenomenon of egress rate does not change significantly with the randomization of variables in the model.

The *Random* behavior is visibly the scenario that causes a higher average egress time, with a maximum of 870 seconds (14.5 minutes) on the maximum capacity scenario in *Alameda*. This can be explained by the sub-optimal exit choice, as a narrow exit has the same probability of choice as a wide exit, causing high levels of congestion in narrow exits. The *Weighted Random* behavior improves the egress flows, proving a balancing of the exit choice by its width. The *Closest* behavior is the scenario with a faster overall egress time, as it decreases the crossing of agents choosing opposing exits.

Comparing the maximum capacity scenarios, we can observe the *Marquês de Pombal* square has the fastest overall egress of 90% of the pedestrians. This could be explained by the largest total exit width it possesses and circular spatial configuration, which allows for a higher egress rate. The egress times of *Alameda* and *Comércio* are close to each other in the random scenarios, even though the latter space has a greater total exit width. This could be explained by the rectangular topology of *Alameda*, which can lead to greater travel distance and agent crossing chance when agents close to an extremity of the rectangle choose exits in the opposite extremity. In the *Closest* scenario, the exit width is decisive on egress times, as the egress performance in *Comércio* is lower than in *Alameda*.

Although the modeled scenarios in this ABM seem plausible, they are improbable, and it may be possible to find more fitting scenarios, using a mix of the defined exit choosing behaviors, incorporating crowd behavior and preferences in the model.

The results of the simulation and the results of the derived theoretical model in [10] have considerable divergences. Upon further analysis of the derived theoretical model, we conclude it needs revision, due to possible erroneous assumptions on egress flow in these scenarios, and will not use it for model validation.

In a normal, non-urgent egress scenario of mass events, it may be hard to predict the crowd movement. The chosen exit would depend greatly on the desired goal. A demographic characterization and observations of egress choices in these type of mass crowd events could help develop more accurate scenarios.

4.5.1 Performance

All simulations were deployed and run on a VM created on top of OpenStack, a free, open standard cloud computing platform used by Infraestrutura Nacional de Computação Distribuída (INCD) to provide computing and storage services to the Portuguese scientific community [4]. The machine runs as operating system the Linux distribution Ubuntu 20-04, VM having 8 virtual CPU and 128 GB of RAM, allowing the running of a suitable number of parallel simulations with decent memory allocated to each.

Although the GAMA platform supports multithreading to model parallel agent routines, the current version of the Pedestrian skill does not support it. Performing sequential SFM calculations for hundreds of thousands of pedestrian agents is time-consuming, making the simulations run slow with that number of agents, and difficult to run in real-time.

[4] https://www.incd.pt/?p=servicos/cloud&lang=en

The 180 simulations were run in batches using a Bash script, which ran batches of 15 parallel GAMA headless instances at a time, each one allocated with a maximum of 8 GB of memory. Each instance ran two sequential scenarios of crowd density for a given square and exit choice behavior.

The heaviest simulation scenarios, with agent numbers reaching around 100,000, took up nearly 6 hours to complete in this environment. Although, execution time of the simulation varies significantly accordingly to the run scenario.

4.5.2 Validity threats

The GAMA version used for the modelling and simulation of this model is in alpha version, meaning the presence of some instability and non-matured functionalities, which could affect model realism and results.

Using the default SFM parameters provided by GAMA may not be optimal to represent the real scenarios represented by the model. A SFM parameter calibration study in non-urgent mass crowd egress conditions should be performed.

More simulations should have been run for each scenario, in order to evaluate the statistical significance of the possible results of the scenarios.

[This page has been intentionally left blank]

Chapter 5.

Urban Fabric Tourism Stress

Contents

This chapter presents the motivation, related work, and the agent-based model to simulate and assess urban fabric crowding stress and results.

[This page has been intentionally left blank]

5.1 Motivation

Many urban environments have high tourism activity. Lisbon is a hotspot for urban tourism, specially the Downtown section and historical neighbourhood of Alfama, from the Santa Maria Maior district.

Historical urban neighbourhoods can have complex, irregular street fabrics, with uneven street configurations. A limiting street fabric could suffer pedestrian bottlenecks from visitors or be problematic for evacuation in case of dangerous events.

An excessive number of visitors to these destinations could be specially damage-inducing, as the old edification networks may not be optimized to accommodate high visitor numbers. There are situations where an abnormal arrival of tourists can come to a city. One of them is the arrival of cruise ships.

Cruise ships can carry a thousands of tourists, and when they arrive to a destination, that amount of tourists flood the venues. If the destination cannot handle the excessive activity derived from cruise tourism, multi-factor damage can occur.

To understand how an urban fabric can accommodate tourists. it is useful to determine its TCC and evaluate the normal tourist behavior when visiting an urban destination. We developed an exploratory ABM to simulate the mass arrival of cruise ship tourists in the Alfama neighborhood, and evaluate the open space pedestrian usage stress in the zone, through the modelling and simulation of tourist walking behavior, evaluating the LOS of street segments.

5.2 Related work

In this section, we present a literature review to analyse relevant state-of-the-art work done on public space egress simulation, following the previously presented Rapid Review methodology.

5.2.1 Research objectives

The objectives of this literature review are:

- Undertake a rapid systematic review of research on spatially explicit tourism activity ABM;
- Filter the articles by relevancy to the given topics to further review in more detail;
- Integrate information from these studies;
- Determine how the various topics approached in the reviews relate to each other through taxonomic classification;
- Identify opportunities for improvement/further investigation.

Concretely, we want to determine:

1. What is the current body of knowledge about the approaches on spatially explicit tourism activity ABM ?

2. What are currently the most effective approaches used for spatially explicit tourism activity ABM?

5.2.2 Review methodology

5.2.2.1 Study selection criteria

Selected articles must present empirical data or the production of publicly available, usable and testable artifacts. The inclusion of studies will not be restricted to any specific type of intervention, but real case scenarios are preferred.

5.2.2.2 Inclusion and exclusion criteria

To include or exclude found articles in this review, the defined criteria are in Tables 5.1 and 5.2

Table 5.1: Inclusion criteria

Inclusion criterion	Description
IC1	Studies must be published in 2011
IC2	Articles must be related to the assessment of PCC of touristic places of interest
IC3	Articles must be research-based
IC4	Studies must be peer-reviewed

Table 5.2: Exclusion criteria

Exclusion criterion	Description
EC1	Articles are inaccessible
EC2	Articles are not written in English

5.2.2.3 Search strategy

The adopted strategy was to search related work on the *Scopus* scientific database, performing the forward and backward snowballing procedure [61].

5.2.2.4 Search strings

Search strings were created through trial and error, combining multiple terms related to the domains of each study, as explained on the previous literature reviews. The search strings used to search on the database are the following:

- TITLE-ABS-KEY (("agent-based"OR "multi-agent"OR "individual-based") AND (simulation OR model*) AND (pedestrian OR visitor OR tourist) AND tourism AND NOT vehicle AND NOT car AND ("city"OR "urban") AND "carrying capacity"AND spatial) AND PUBYEAR > 2010

5.2.2.5 Selection process

The selection process steps are defined as:

1. Execute the search string on the defined databases and search engines;
2. Apply the inclusion criteria (orderly by title, abstract, keywords and content);
3. Apply the exclusion criteria;
4. Apply backward and forward snowballing to the remaining articles;
5. Repeat steps 2 and 3 on the new articles selected by snowballing.

5.2.2.6 Validity threats

This review has possible threats to its validity. First, the number of articles used for the analysis of the related work may be less than ideal to assure the validity of the results. Out of all the related work found, only a total of 4 articles were selected.

First, only using a single search engine may affect validity, as there may be useful related studies which have not been properly indexed, missing on them.

Second, more variations of the search strings, such as appending more terms and varying term combinations, could have been used to narrow the database results to a lesser number of more relevant articles. Limiting the search strings might also have led to not finding some related work.

Last, inaccessible articles were discarded, which could lead to the loss of important information. An effort to reach the authors should have been done.

5.2.3 Classification taxonomy

5.2.3.1 Taxonomy criteria

To categorize the different types of spatially-explicit tourist activity simulation, a set of criteria was defined. The Table 5.3 shows the selected criteria.

Table 5.3: Taxonomy criteria

Criterion	Acronym
Scenario realism	SR
Modelling scale	MS
Space type	ST
Validation	V

5.2.3.2 Scenario Realism (SR)

The selected studies can have egress simulations to describe real-life places through static digital twinning, or simply study pedestrian dynamics in fabricated scenarios. This criterion categorizes the simulation scenario realism by the following ordinal scale:

- **Abstract Environment**(AE): The modeled environment is fabricated, and does not represent a real-life environment;

- **Real Environment**(RE): The modeled environment in simulations represents a real-life location.

5.2.3.3 Modelling Scale (MS)

Modelling can be divided into two scales: microscopic and microscopic. A microscopic model considers the interaction of individual agents, while a macroscopic model considers a higher level collective behavior of multiple entities. The criterion categorizes the models in the studies by the following nominal scale:

- **Microscopic (MIC)**: The model is microscopic;
- **Macroscopic (MAC)**: The model is microscopic.

5.2.3.4 Space Type (ST)

Egress models can describe indoor or outdoor scenarios. This criterion categorizes by the following nominal scale:

- **Open space (OS)**: The modeled environment represents an open, outdoor space, such as a city square.
- **Enclosed space (ES)**: The modeled environment in simulations represents enclosed environments, such as rooms in buildings, train stations, etc.

5.2.3.5 Validation (V)

This criterion describes if a study is validated or not. The following ordinal scale categorizes the criterion:

- **Not Validated (NV)**: The study is not validated.
- **Validated (V)**: The study is validated.

5.2.4 Related work review

5.2.4.1 Bolshakov et al.,*"Simulation in Intelligent Management of Pedestrian Flows at Heritage Sites"*, 2016 [7]

Objective: The main objective of the present study is to present an ABM to simulate the management of pedestrian flows with a mobile application that informs its users on the number of visitors in tourist attractions, namely heritage sites, suggesting the visiting of less crowded attractions.

Technical Summary: The model was implemented in the AnyLogic platform [1], using its Pedestrian library. The modeled environment is fabricated, not corresponding to a real location. It

[1] https://www.anylogic.com/

consists on an area that appears urban, containing buildings, pedestrian and vehicular streets and indoor and outdoor attraction places with limited capacity.

The tourist agents either use or do not use a recommending software that allows avoidance of overcrowded attractions, choosing randomly the closest attractions. When all attractions are visited, the tourist agents leave the environment.

A set of simulations were performed, where the proportions of tourists with and without the application are changed. The results shown that increasing the percentage of informed tourists to 50% decreases queuing in attraction, but reaching to 100% would not improve the scenario. Thus, the authors conclude the model and the mobile application concept should be improved.

As future research work, Bolshakov et. al point the model should be improved with more advanced agent logic, the inclusion of factors such as time, real distance to attractions, walked distance and overall satisfaction of visitors. The simulation model could represent real-life locations with geospatial data, and be used to assess bottlenecks.

Classification:

Table 5.4: Classification of *"Simulation in Intelligent Management of Pedestrian Flows at Heritage Sites"*

Source	SR	MS	ST	V
[7]	AE	MIC	ES, OS	1

5.2.4.2 Grignard et al.,*"CityScope Andorra data observatory: A case study on tourism patterns"*, 2019 [27]

Objective: The main objective of the present study is to present a data-driven ABM to simulate individual mobility of tourists based on mobile phone spatio-temporal data in Andorra.

Technical Summary: The presented model is integrated in the CityScope framework, a system to aid in urban planning. It was implemented in the *Processing* tool [2] and the GAMA platform [63]. The ABM describes and helps understand the visitor flows of Andorra and traffic congestion during 2 big events: *Cirque du Soleil: VISION* and *La Tour de France*.

The model uses Mobile data, Amenities data and Road network data from OSM to describe the model and agent behavior. The agents, representing tourist pedestrians and vehicles, move according to the capacity of amenities, working hours, events, road direction, traffic congestion and amenity occupancy in simulation real-time.

The visualization of the simulation results shows different patterns of movements from visitors, revealing the city complex system. This model allowed to understand key differences in the tourist activity generated by the 2 events.

As future work, the authors suggest the usage of a new source of data from *Andorra Telecom* that can provide the geolocation of the devices, the integration of sensor data in stores to monitor indoor behavior, and the replay of agents' behavior to understand city dynamics and

[2] https://processing.org/

improve urban designs.

Classification:

Table 5.5: Classification of *"CityScope Andorra data observatory: A case study on tourism patterns"*

Source	SR	MS	ST	V
[27]	RE	MIC	ES, OS	1

5.2.4.3 Zhang et al.,*"Coupling Social Media and Agent-Based Modelling: A Novel Approach for Supporting Smart Tourism Planning"*,2020 [72]

Objective: The main goal of the present study is to present a social-media-backed ABM to simulate tourist movement and decision-making, analyzing tourist preferences. The model is applied in the Zaolinwan Park, China.

Technical Summary: Social media and questionnaire data is used to identify the preferences of potential visitors to the Zaolinwan Park, such as preferred activities, food, and accommodation choices, defining tourist typologies. The identified tourist behaviors and preferences are then used to develop an ABM to simulate the decision-making process of tourists.

The visitation patterns of agents are randomized, but consider destination choices, activity/food/accommodation preferences and transportation factors.

The spatio-temporal patterns of tourist distribution and the interactions among different functional zones, shown by the simulation results, could aid the planning facilities and service provision in the tourism park.

As future work, Zhang et al. suggest further analysis of on how social-media-based practices can fit into formal decision-making procedures, and the interviewing of the perspectives of authorities to promote the presented approach and evaluate its potential. The authors also note the method could be applied in the planning of other infrastructures and services at community and urban levels.

Classification:

Table 5.6: Classification of *"Coupling Social Media and Agent-Based Modelling: A Novel Approach for Supporting Smart Tourism Planning"*

Source	SR	MS	ST	V
[72]	RE	MIC	OS	1

5.2.4.4 Bai et al.,*"Towards a finer heritage management: Evaluating the tourism carrying capacity using an agent-based model"*, 2020 [2]

Objective: The study aims to present an audio-guided touring program based on the dynamic evaluation of the TCC to optimize tours, and an ABM to simulate the performance of the program in the Hall of Mental Cultivation from the Palace Museum of Beijing.

Technical Summary: The proposed touring program calculates a sequence of locations and the ideal stay time in each, using a defined tourist arrival lag, and dynamically calculated values for maximum and suitable TCC for each location according to its type (outdoor, roofed, indoor) and importance.

The authors developed an ABM to simulate the performance of the program, based on the Thunderhead Pathfinder platform [3], where each tourist agent follows the recommended route, but has randomly assigned times in each spot with a Normal distribution, and has different freedoms according to space type.

Through a set of simulations and parameter adjustments, the performance of the touring program in all spots was kept reasonable at all times, meeting all the objectives, which are not surpassing the TCC at any time in any spot; 2) controlling and smoothing peak of the tourist arriving periodicity along the tour and avoid long queues in checkpoints.

It is concluded the proposed touring program allows to improve visiting experience, heritage safety, and management efficiency. Although the model may not accurately describe real conditions due to unconsidered attributes, it is enough to give insight on the exhibition planning and management program.

As future work, the authors pretend to use tracking tools in the Hall for activity monitoring and validate/calibrate the model.

Classification:

Table 5.7: Classification of *"Towards a finer heritage management: Evaluating the tourism carrying capacity using an agent-based model"*

Source	**SR**	**MS**	**ST**	**V**
[2]	RE	MIC	ES	1

Table 5.8: Summary table with the related work classification

Source	**SR**	**MS**	**ST**	**V**
[7]	AE	MIC	ES, OS	1
[27]	RE	MIC	ES, OS	1
[72]	RE	MIC	OS	1
[2]	RE	MIC	ES	1

5.2.5 Conclusions

The spatially explicit tourist behavior models can range from inter-region flows, down to the flows inside museums. In this review, we focus on the studies with lower spatial scope, up to city scale. Most of models in the reviewed studies simulated real locations, and the ones using fabricated environments could be adapted to the real destinations. Every reviewed study used microscopic modelling, which seems fit to simulate urban tourism usage, but a macroscopic model also seems adequate. The tourism behavior is usually modeled using multiple data

[3] https://www.thunderheadeng.com/pathfinder

sources and assumptions. There are spatial models addressing TCC, but address building scale scenarios. A gap in the literature for historical neighborhood tourism visitation capture and simulation is evident.

5.3 Agent-based model

This ABM is designed to simulate and evaluate pedestrian usage and crowding stress of a historical neighbourhood close to a cruise ship terminal, during the arrival of tourists from cruise ships. The agents will model tourists and local residents, who also use the urban space, navigating in the neighborhood, simulating its usage. Tourism places of interest from the neighborhood are represented in the environment as tourism activity hotspots, which the tourists want to visit. The PCC assessing open space segments created by the algorithm presented in 3 serve as crowd monitoring agents, storing the crowding levels of service throughout the simulations.

5.3.1 Simulation software

As previously stated in the dissertation, the used ABM platform is the GAMA 1.8.2 alpha build.

5.3.2 Environment modelling

The model environment space envelopes all Alfama neighborhood and the eastern Santa Maria Maior area, bordering the Lisbon Cruise Terminal. The model environment consists of building geometries, and a list of georeferenced points representing highly visited and well known tourism places of interest on the covered zone. The places represented in this model are the area in front of the entrance of the *Sé* church, the entrance to *São Jorge* castle, the *Santa Luzia* lookout and *Portas do Sol* lookout.

This data was downloaded from OSM and processed using *QGIS* 3.16 LTS, creating ESRI shapefiles to be read into the model. The usable open area of the urban fabric is represented by the open space segments obtained through the algorithm described in Chapter 3. A graph to allow the pedestrians to navigate the open space of the zone, using various types of graph routing algorithms, was automatically built with primitives from the Pedestrian plugin included in GAMA. In Figure 5.1 is shown a satellite image of the modeled zone, which is inside the red borders. The image was taken from the *geojson.io*[4] tool, using satellite tile layers from Maxar [5]. Note that it is possible to observe a docked cruise ship in the cruise terminal. In Figure 5.2, we can observe the modeled environment, with the building geometries and the points of interest, represented by the red dots.

5.3.3 Agent modelling

In this model, there are 3 types of pedestrian agents: The local residents, the land tourists and cruise arriving tourists. The local residents represent the population that lives in the neighbourhood. They are modeled as having totally unpredictable, random behavior, not

[4] https://geojson.io/

[5] https://www.maxar.com/

Figure 5.1: Satellite image of the Alfama neighbourhood and eastern Santa Maria Maior, with the boundaries of the modeled area in red (source: geojson.io, Maxar)

Figure 5.2: Snapshot of the modeled environment of the Alfama neighbourhood and eastern Santa Maria Maior in GAMA

having particular interest in visiting tourist destinations. The land tourists are tourists who are already visiting the city when the cruise ships arrive. These tourists may also be locals, but may have interest in visiting touristic places of interest. The cruise tourists are tourists who arrive from cruises. These tourists will have similar destination choosing behavior as land tourists, but will spawn in front of the the Lisbon Cruise Terminal, near the doors. The behavior of tourists is inspired by the one modeled in [7]. The urban fabric open space segments are also active agents, counting the number of pedestrians inside the segments, monitoring and saving the crowding LOS throughout the simulation.

5.3.3.1 Agent behavior

The model agents behave as follows:
Cruise tourist: Cruise tourist agents spawn near the doors of the Lisbon Cruise Terminal, one every 2 seconds in simulation time, simulating an orderly arrival of the tourists to the city. When they spawn, they randomly choose one of points of interest in the model, weighted on the inverse of the distance, i.e., the closer the destination, the more probable is its to be chosen. At each simulation cycle, the agents behave as follows:

1. Compute the shortest path to its desired destination, using the generated pedestrian path graph;

2. The agent calculates the sum of the social forces acting upon it, and the desired velocity force. If obstacles and/or pedestrians are in the consideration distance to the agent, they will exert social forces. The final speed and direction of the agent are calculated;

3. The agent moves according to the calculated speed and direction;

4. If the agent has not visited all available destinations, a new random destination will be chosen, using the weight on the inverse of the distance. Else, it will choose an uniformly random location on the open space, behaving as a local.

Land tourist: Land tourists are spawned in uniform distribution in random locations in the open space. Land tourists have equal place of interest choosing behavior to cruise tourists.
Local Resident: The locals are spawned in uniform distribution in random locations in the open space, similar to land tourists. In the beginning to the simulation, an agent will choose a random location as its destination in the model open space.

During each cycle, the agents follow these instructions:

1. Compute the shortest path to its desired destination, using the generated pedestrian path graph;

2. The agent calculates the sum of the social forces acting upon it, and the desired velocity force. If obstacles and/or pedestrians are in the consideration distance to the agent, they will exert social forces. The final speed and direction of the agent are calculated;

3. The agent moves according to the calculated speed and direction;

4. If the agent reaches its chosen destination, a new random destination will be chosen.

Open Space Segment: At each cycle, each segment counts the number of pedestrian agents that are contained inside its geometry. Having the PCC for each pedestrian LOS of a segment, it will evaluate which is the pedestrian LOS at the time cycle. Depending on the LOS, a variable representing the total cycles in which the specific LOS was detected is increased by 1, allowing to store the total segment pedestrian usage during a simulation. Figure 5.3 shows a snapshot of a locally run simulation, showing the *inspect* functionality from GAMA, demonstrating the variables and respective values of a segment agent.

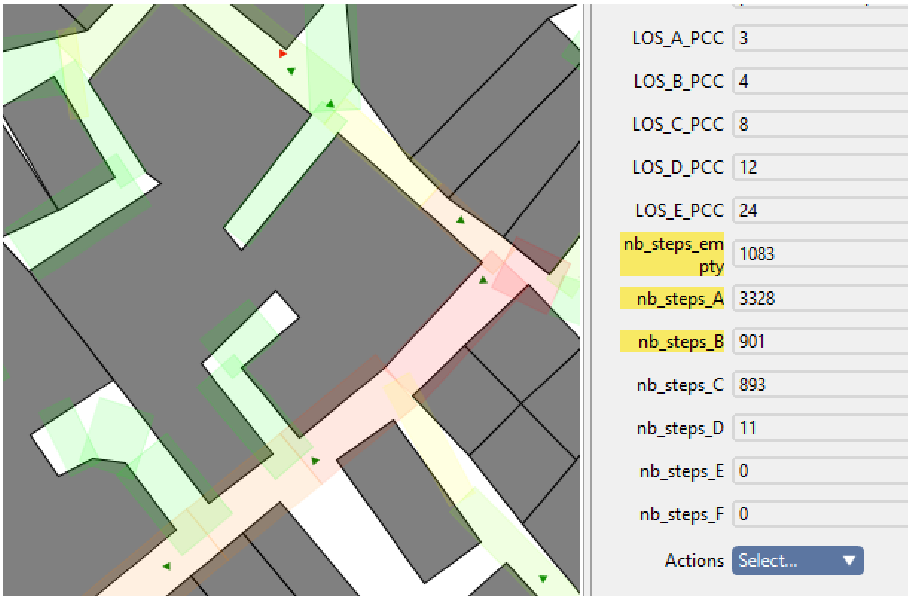

Figure 5.3: Snapshot of the *inspect* function of a segment agent, showing its PCC values and LOS counters

5.3.3.2 Pedestrian skill parameters

The values for the parameters for the Pedestrian skill are the default "classic"values in GAMA. In Table 5.9 are displayed the available parameters for the skill and the set values for said parameters.

Table 5.9: Pedestrian skill SFM parameters, the description of the parameters from the GAMA documentation and value

SFM Parameter	Description from GAMA documentation	Value
shoulder_length	The width of the pedestrian (in meters) - classic values: [0.39, 0.515]	0.45
A_pedestrians_SFM	Value of A in the SFM model for pedestrians - the force of repulsive interactions (classic values : mean = 4.5, std = 0.3)	4.5
A_obstacles_SFM	Value of A in the SFM model for obstacles - the force of repulsive interactions (classic values : mean = 4.5, std = 0.3)	4.5
B_pedestrians_SFM	Value of B in the SFM model for pedestrians - the range (in meters) of repulsive interactions	2.0
B_obstacles_SFM	Value of B in the SFM model for obstacles - the range (in meters) of repulsive interactions	2.0
k_SFM	Value of k in the SFM model: force counteracting body compression	200
kappa_SFM	Value of kappa in the SFM model: friction counteracting body compression	400
relaxion_SFM	Value of relaxion in the SFM model - the amount of delay time for an agent to adapt .(classic values : mean = 0.54, std = 0.05)	0.54
gama_SFM	Value of gama in the SFM model the amount of normal social force added in tangential direction. between 0.0 and 1.0 (classic values : mean = 0.35, std = 0.01)	0.35
lambda_SFM	Value of lambda in the SFM model - the (an-)isotropy (between 0.0 and 1.0)	0.5
n_SFM	Value of n in the SFM model (classic values : mean = 2.0, std = 0.1)	2.0
n_prime_SFM	Value of nín the SFM model (classic values : mean = 3.0, std = 0.7)	3.0

5.3.4 General assumptions

The general assumptions in the basis of the models are:

- Pedestrians are created with walking speeds following normal distribution with an average of 1.34 m/s and a standard deviation of 0.26 m/s [30]
- Group behavior is neglected. Every agent is considered alone, and has individual behavior;

- All area without buildings is considered usable area for the pedestrian agents. Although this facilitates pedestrian walking, the PCC assessment considers estimated road area.
- Altitude is not considered in the model for simplicity. The entire model area is considered a two-dimensional plane.

5.4 Simulation scenarios

The model will be simulated in a total of 5 scenarios. Each scenario is solely defined by the number of tourists and locals. The number of cruise tourists is set at 4,500 pedestrians, indicated as the maximum number of passengers the terminal can handle at a given time [52]. In each variation, the number of land tourists and locals increase each by 500, starting at 500 people each. In the last scenario, a total of 9,500 agents will traverse the model environment. A simulation will have a model duration of 8 hours, or 28,800 seconds. The simulation step for each cycle is 1 second, totaling 28,000 cycles. As output of each simulation, ESRI shapefiles with the open space segments and respective pedestrian usage at the end of the simulations and a snapshot of the simulation visualization for each simulation step.

5.5 Results

In Figure 5.4, we can observe the results of the scenario with 9,500 pedestrian agents, with the open space segments, colored with a green-yellow-red scale, according to the percentage of time the segments were overcrowded. The coloring is relative to the maximum percentage value in a segment. In this scenario, the maximum segment with 0.035% of overcrowding time, corresponding to overcrowding time of 16.8 minutes. The dark red painted segments are the segments with the most crowding stress, and the green painted ones are the least used segments. Using as overcrowding threshold the PCC of LOS C, any second the segment has higher pedestrian usage than this limit is considered overcrowded time.

As can be observed in Figure 5.4, the segments near the points of interest have the red color, indicating a high average usage during the simulation time. Since there are few destinations represented, and every location is visited by the tourists, an elevated usage is expected. Other segments unrelated to the tourism hotspots show higher overcrowding percentages, possibly due to the modeled random behavior. Narrow streets that connect the interest locations also show higher usage, naturally due to their smaller sizes.

The current iteration of model simulation only stores the number of seconds each segment has a given LOS, but it can store crowding information for each time step, allowing to perform more complex spatio-temporal analysis. The random behaviors could impact the spatio-temporal distribution of pedestrian usage.

Even though the model lacks some richness in terms of the representation of tourism hotspots and pedestrian (tourist) behavior, the results show the feasibility of the model to simulate and capture crowding stress/pedestrian usage given a tourism scenario, which is in this case, the arrival of cruise ship tourists.

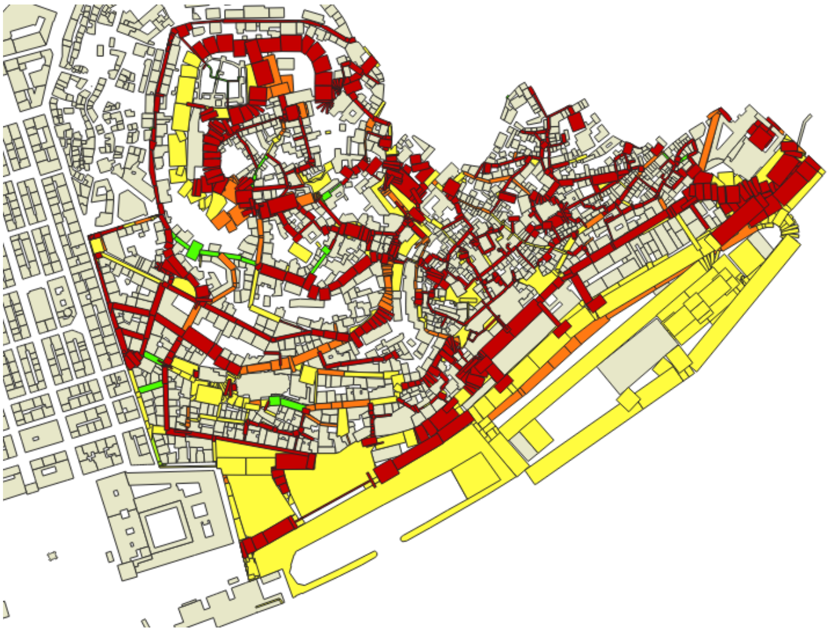

Figure 5.4: Results of the heaviest simulation scenario. The segments are colored from green to red to yellow by overcrowding time percentage

5.5.1 Performance

The environment where the simulations of the model run are the same as the egress model presented in Chapter 4. A total of 5 simulations were run, one of each scenario. A Bash script run 5 parallel instances of GAMA headless, each running a single simulation. This model is more resource-demanding than the previously described egress model. Although the number of active agents is considerably less in this model, given the heaviest scenarios of both models, the execution time is considerably longer in this model. The heaviest scenario, with 9,500 pedestrian agents, took aproximately 2 days and 2 hours to complete 8 hours of simulation time.

The complexity of the Pedestrian skill walking features may not add value to the present model, as an inter-pedestrian interaction model with such a small scale may be too microscopic for this scenario, and simpler pedestrian routing mechanisms could improve simulation speed.

5.6 Validity threats

The GAMA version used for the modelling and simulation of this model is in alpha version, meaning possible instability and non-matured functionalities, which could affect model realism. Graph open space do not cover the entirety of the available open space, which limit the open space usage of the pedestrian agents, and creates paths which are not normally accessible in the real environment, possibly leading to less realistic routing. The model may be too simple to accurately describe the emergent phenomena detected in the simulations. Further validation and model improvement are needed.

[This page has been intentionally left blank]

CHAPTER

6.

CONCLUSION

Contents

In this chapter, we present the outcome of the studies performed in this dissertation, as well as future work directions.

[This page has been intentionally left blank]

6.1 Summary

To mitigate the problem of Overtourism, real-time tourism control measures, based on real-time or predicted crowding measurements, could be applied. These types of measures depend on the understanding of tourism crowding dynamics, and the effective measurement and classification of crowding states.

To measure and classify localized crowding states, the PCC is a fit initial value. In order to obtain said value for different locations throughout an urban fabric, an algorithm to segment and assess PCC was developed.

To gain insight on artificial and organic tourism crowding scenarios, two ABM were developed. The scenarios described by each model are (i) the visitor non-urgent egress from mass crowd events and (ii) the abnormal tourism visitation of a constrained urban fabric.

Although the ABM describe different scenarios, there are some similarities in terms of modelling, such as the individual modeling of pedestrians via agents who interact with each other through SFM.

In the urban fabric visitation ABM, the PCC segments generated by the presented algorithm act as monitoring agents, capturing their own crowding states, allowing to assess the feasibility of the algorithm in crowd monitoring.

6.2 Conclusions

In this dissertation we proposed an algorithm implementation for creating urban fabric geospatial segments for area estimation and PCC assessment using OSM data; an ABM for simulating non-urgent mass crowd event egress and an ABM to simulate the spatial touristic usage of a historical neighborhood during the arrival of tourists from cruise ships.

The algorithm for PCC assessment and open space segmentation showed acceptable results, capturing most of the urban open space and representing it with enough accuracy for engineering purposes. The incompleteness of OSM data brought the need to estimate usable areas in urban spaces using a set of assumptions based on observations and street design information from Lisbon's Public Space Manual [12].

The ABM for simulating non-urgent mass crowd event egress was able to simulate the egress of crowded events from 3 Lisbon squares, in different crowd densities and exit choosing behaviors. Variables such as crowd density, total exit width and the spatial topology of the open spaces have individual impacts on the total egress times. *Marquês do Pombal* has the fastest overall egress time, and *Alameda* and *Comércio* can have superior egress performances to each other depending on the exit choice behavior.

The ABM for simulating historical urban neighbourhood crowding stress, even though in an initial stage of development, showed the feasibility to capture urban open space usage using the geospatial segments created with the previously described algorithm, given modeled tourist behaviors. Further improvement of the tourist behavior and environment modelling is required, as well as validation.

Agent-based simulations with a large number of agents and complex behaviors are CPU and memory intensive, being hard to run on machines with lower specifications in reasonable time. The execution time of certain models may be reduced by harnessing parallel processing in the models.

6.3 Future work

For each study in this dissertation, we present future work directions and suggest further improvements to the current work:

Urban fabric segmentation and PCC assessing algorithm: The area determination algorithm should be improved for greater accuracy. Satellite image analysis with Machine Learning techniques or more spatially fine-grained and precise algorithms using vectorial geographic data could be used to improve the algorithm. Other data sources and in-field measurements could be used to further improve accuracy of total and usable area estimation, and the validation of the algorithm. Testing the accuracy on a greater number of streets and open places in the Santa Maria Maior district, as well as other locations, should be performed for a more solid validation of the algorithm. Since the algorithm relies on the accuracy of the OSM data, the improvement of that data is important. Contributing to the mapping project with volunteered mapping is a solid solution for the improvement of data coverage and quality;

Although the PCC is a useful measure for engineering purposes, it may not represent the actual CC of the urban destination. The ECC and RCC should be assessed, following an adapted Cifuentes methodology for urban places, determining the adequate corrective factors and management capacity of the destination;

Open space mass event non-urgent egress ABM: To improve model realism, observation and studies of mass crowd non-urgent egress conditions should be performed, to calibrate SFM parameters to more fitting values and to detect macroscopic and microscopic behavior to be modeled in the agents.

The model should compared with implementations with other modelling techniques, such as cellular automata, macroscopic modelling or optimized versions of SFM, in order to find the most suitable model technique for this scenario.

The used GAMA alpha build has some stability issues regarding the Pedestrian skill and other functionalities (as expected from an alpha build). The model should be ported to a stable version in the future.

Urban fabric usage simulation ABM: The current modeled pedestrian agent behavior is quite simple, which may not accurately represent real pedestrian urban fabric usage. As such, it is needed to improve pedestrian agent intelligence for decision making;

Studies to understand the real life Alfama's spatial and touristic usage by pedestrian type should be performed to gain insight about the tourist behavior and more accurately model their agent digital twins. Such data could be obtained by flow study from mobile operators, census data,local accommodations and social media.

A more complete environment should also be modeled, with the inclusion of more tourism places of interest and real data about their attractiveness.

As previously explained, the used GAMA alpha build has some stability issues, thus this model should also be ported to a stable version.

Bibliography

[1] B. A. Adie, M. Falk, and M. Savioli. "Overtourism as a perceived threat to cultural heritage in Europe." In: *Current Issues in Tourism* 23.14 (2020), pp. 1737–1741.

[2] N. Bai, P. Nourian, A. Xie, and A. Pereira Roders. "Towards a Finer Heritage Management-Evaluating the Tourism Carrying Capacity using an Agent-Based Model." In: (2020).

[3] R. Beirão. "A Multiple Criteria Route Recommendation System." Master's thesis. Iscte – Instituto Universitário de Lisboa, 2020.

[4] D. Bertocchi, N. Camatti, and J. van der Borg. "Tourism Peaks on the Three Peaks. Using big data to monitor where, when and how many visitors impact the Dolomites UNESCO World Heritage Site." In: *Rivista geografica italiana* 110.3 (2021), pp. 59–81.

[5] L. Blaga and I. Josan. "Estimating the tourist carrying capacity for protected areas. A case study for Natura 2000 sites from NorthWestern Romania." In: *Forum Geografic* 18.1 (2019), pp. 71–82. doi: 10.5775/fg.2019.067.i.

[6] I. Boavida-Portugal, C. C. Ferreira, and J. Rocha. "Where to vacation? An agent-based approach to modelling tourist decision-making process." In: *Current Issues in Tourism* 20.15 (2015), pp. 1557–1574. doi: 10.1080/13683500.2015.1041880.

[7] V. Bolshakov and G. Merkuryeva. "Simulation in Intelligent Management of Pedestrian Flows at Heritage Sites." In: *2016 International Conference on Systems Informatics, Modelling and Simulation (SIMS)*. IEEE. 2016, pp. 18–22.

[8] E. Bonabeau. "Agent-based modeling: Methods and techniques for simulating human systems." In: *Proceedings of the National Academy of Sciences of the United States of America* 99.SUPPL. 3 (2002), pp. 7280–7287. doi: 10.1073/pnas.082080899.

[9] R. C. Boullón. *Planificación del espacio turístico*. Fourth. Trillas, 1985, p. 245.

[10] F. Brito e Abreu and D. Almeida. "Capacidade de Carga e Tempo de Evacuação para Eventos nos Espaços Públicos das Cidades." In: *IntelCities* (2021). doi: 10.5281/ZENODO.4600967.

[11] C. Burstedde, K. Klauck, A. Schadschneider, and J. Zittartz. "Simulation of pedestrian dynamics using a two-dimensional cellular automaton." In: *Physica A: Statistical Mechanics and its Applications* 295.3-4 (2001), pp. 507–525.

[12] Camara Municipal de Lisboa. *Lisboa: O Desenho da Rua: Manual de espaço público*. Lisboa : CML - Departamento de Espaço Público, Departamento de Urbanismo, 2018.

[13] N. Camatti, D. Bertocchi, H. Carić, and J. van der Borg. "A digital response system to mitigate overtourism. The case of Dubrovnik." In: *Journal of Travel & Tourism Marketing* 37.8-9 (2020), pp. 887–901.

[14] B. Cartaxo, G. Pinto, and S. Soares. "Rapid Reviews in Software Engineering." In: *Contemporary Empirical Methods in Software Engineering* (2020), pp. 357–384. DOI: 10.1007/978-3-030-32489-6_13. arXiv: 2003.10006.

[15] H. Ceballos-Lascurain. "Tourism, Ecotourism, and Protected Areas: The State of Nature-Based Tourism around the World and Guidelines for Its Development." In: *The Geographical Journal* 164.3 (1998), p. 349. DOI: 10.2307/3060626.

[16] M. Cifuentes. *Determinación de Capacidad de Carga Turística en Áreas Protegidas*. 1992.

[17] L. Crociani, A. Piazzoni, G. Vizzari, and S. Bandini. "When reactive agents are not enough: Tactical level decisions in pedestrian simulation." In: *Intelligenza Artificiale* 9.2 (2015), pp. 163–177.

[18] L. Crociani, G. Vizzari, D. Yanagisawa, K. Nishinari, and S. Bandini. "Route choice in pedestrian simulation: Design and evaluation of a model based on empirical observations." In: *Intelligenza Artificiale* 10.2 (2016), pp. 163–182.

[19] R. C. De Sousa, L. C. Pereira, R. M. Da Costa, and J. A. Jiménez. "Tourism carrying capacity on estuarine beaches in the Brazilian Amazon region." In: *Journal of Coastal Research* 70.70 (2014), pp. 545–550. DOI: 10.2112/SI70-092.1.

[20] R. Dias da Silva, R. Neto Marinheiro, and F. Brito e Abreu. "Crowding Detection Combining Trace Elements from Heterogeneous Wireless Technologies." In: *22nd International Symposium on Wireless Personal Multimedia Communications (WPMC)*. IEEE. 2019, pp. 1–6.

[21] A. Dichter and G. M. Gloria. *Coping with Success : Managing Overcrowding in Tourism Destinations*. Tech. rep. London, UK: WTTC and McKinsey & Company, 2017, pp. 1–56.

[22] R. W. Douglass. *Forest Recreation*. Pergamon, 1982, p. 336.

[23] N. Dragoni, I. Lanese, S. T. Larsen, M. Mazzara, R. Mustafin, and L. Safina. "Microservices: How to make your application scale." In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 10742 LNCS (2018), pp. 95–104. DOI: 10.1007/978-3-319-74313-4_8. arXiv: 1702.07149.

[24] D. Getz. "Capacity to absorb tourism. Concepts and implications for strategic planning." In: *Annals of Tourism Research* 10.2 (1983), pp. 239–263. DOI: 10.1016/0160-7383(83)90028-2.

[25] T. H. Göktuğ, Y. Bulut, M. Demir, N. D. Yıldız, et al. "Estimating carrying capacity of Olympos Bey Mountains Coastal National Park." In: *Fresenius Environmental Bulletin* 22.12 (2013), pp. 3730–3738.

[26] T. H. Göktuğ, Y. Bulut, N. D. Yildiz, and M. Demir. "Carrying capacity assessment of Tortum Waterfall, Turkey." In: *Fresenius Environmental Bulletin* 22.12 B (2013), pp. 3783–3791.

[27] A Grignard, L Alonso, N Macia, M Vilella, and K Larson. "Cityscope andorra data observatory: A case study on tourism patterns." In: *CEUR Workshop Proceedings. CEUR*. 2019.

[28] M. Hampton and J. Hampton. "Is the beach party over? Tourism and the environment in small islands: A case study of Gili Trawangan, Lombok, Indonesia." In: *Tourism in Southeast Asia: Challenges and new directions* 286 (2009).

[29] D. Helbing. *FuturICT Blog: SOCIAL FORCES: Revealing the causes of success or disaster*. 2014.

[30] D. Helbing and P. Molnár. "Social force model for pedestrian dynamics." In: *Physical Review E* 51.5 (1995), pp. 4282–4286. doi: 10.1103/PhysRevE.51.4282. arXiv: 9805244 [cond-mat].

[31] S. Heliövaara, H. Ehtamo, D. Helbing, and T. Korhonen. "Patient and impatient pedestrians in a spatial game for egress congestion." In: *Physical Review E* 87.1 (2013), p. 012802.

[32] S Holl and A Seyfried. "Laboratory experiments on crowd dynamics." In: *inSiDE (Innovatives Supercomputing in Deutschland)* 11.2 (2013).

[33] M. A. Huamantinco Cisneros, N. V. Revollo Sarmiento, C. A. Delrieux, M. C. Piccolo, and G. M. Perillo. "Beach carrying capacity assessment through image processing tools for coastal management." In: *Ocean and Coastal Management* 130 (2016), pp. 138–147. doi: 10.1016/j.ocecoaman.2016.06.010.

[34] M. J. Hurley, D. T. Gottuk, J. R. Hall Jr, K. Harada, E. D. Kuligowski, M. Puchovsky, J. M. Watts Jr, C. J. WIECZOREK, et al. *SFPE handbook of fire protection engineering*. Springer, 2015.

[35] F. T. Illiyas, S. K. Mani, and N. Babu. "Carrying capacity assessment for religious crowd management - An application to sabarimala mass gathering pilgrimage, India." In: *International Journal of Religious Tourism and Pilgrimage* 8.8 (2021), pp. 79–93. doi: 10.21427/82RW-FX96.

[36] R. M. Itami. "Estimating Capacities for Pedestrian Walkways and Viewing Platforms." In: *A report to Parks Victoria 20* June (2002).

[37] R. Jangra and S. P. Kaushik. "Assessment of physical carrying capacity for managing sustainability at religious tourist destinations." In: *International Journal of Religious Tourism and Pilgrimage* 5.1 (2017), pp. 29–38. doi: 10.21427/D72X4N.

[38] J. Kennell. "Carrying capacity, tourism." In: *Encyclopedia of Tourism* 1983 (2014). doi: 10.1007/978-3-319-01669-6.

[39] K. Koens, A. Postma, and B. Papp. "Management strategies for overtourism: From adaptation to system change." In: *Overtourism*. Routledge, 2019, pp. 149–160.

[40] F. Larga. "A Mobile Tour Guide App for Sustainable Tourism." Master's thesis. Iscte – Instituto Universitário de Lisboa, 2020.

[41] K. Lewin. "Field theory in social science: selected theoretical papers." In: *The Annals of the American Academy of Political and Social Science* (1951). Ed. by D. Cartwright. doi: 10.1177/000271625127600135.

[42] C. M. Macal. "Everything you need to know about agent-based modelling and simulation." In: *Journal of Simulation* 10.2 (2016), pp. 144–156. DOI: 10.1057/jos.2016.7.

[43] A. M. Madni, C. C. Madni, and S. D. Lucero. "Leveraging digital twin technology in model-based systems engineering." In: *Systems* 7.1 (2019), p. 7.

[44] A. Makhadmeh, M. Al-Badarneh, A. Rawashdeh, and A. Al-Shorman. "Evaluating the carrying capacity at the archaeological site of Jerash (Gerasa) using mathematical GIS modeling." In: *Egyptian Journal of Remote Sensing and Space Science* 23.2 (2020), pp. 159–165. DOI: 10.1016/j.ejrs.2018.09.002.

[45] S. F. Mc Cool and D. W. Lime. "Tourism carrying capacity: Tempting fantasy or useful reality?" In: *Journal of Sustainable Tourism* 9.5 (2001), pp. 372–388. DOI: 10.1080/09669580108667409.

[46] A. Micallef and A. T. Williams. "Application of a novel approach to beach classification in the Maltese Islands." In: *Ocean and Coastal Management* 47.5-6 (2004), pp. 225–242. DOI: 10.1016/j.ocecoaman.2004.04.004.

[47] G. G. Morales, J. A. Arreola-Lizárraga, and P. R. Grano. "Integrated Assessment of Recreational Quality and Carrying Capacity of an Urban Beach." In: *Coastal Management* 46.4 (2018), pp. 316–333. DOI: 10.1080/08920753.2018.1474070.

[48] I. Moukhtar, A. A. Ibrahim, T. A. El Seoud, and S. Mostafa. "Assessment of the environmental carrying capacity for protected areas: A study of petrified forest and hassanah dome, the great cairo." In: *Environmental Research, Engineering and Management* 76.4 (2020), pp. 106–117. DOI: 10.5755/j01.erem.76.4.26665.

[49] Norma-Cubana. *Sistema de Normas para la protección del Medio Ambiente: áreas de playa*. Tech. rep. Comité Estatal de Normalización, 1988.

[50] W. T. Organization. *Technical Handbook on the Collection and Presentation of Domestic and International Tourism Statistics: World Tourism Organization*. World Tourism Organization, 1981.

[51] R. Peixoto. "A graph-based approach for sustainable walking tour recommendations: the case of Lisbon overcrowding." Master's thesis. Iscte – Instituto Universitário de Lisboa, 2019.

[52] Público. *Novo terminal de cruzeiros de Lisboa inaugurado hoje, meses depois do previsto*. 2017.

[53] I. Rodella, C. Corbau, U. Simeoni, and K. Utizi. "Assessment of the relationship between geomorphological evolution, carrying capacity and users' perception: Case studies in Emilia-Romagna (Italy)." In: *Tourism Management* 59 (2017), pp. 7–22. DOI: 10.1016/j.tourman.2016.07.009.

[54] N. M. Rouphail and D. P. Allen. "Recommended Procedures Chapter 13, "Pedestrians,"of the Highway Capacity Manual." In: *Highway Capacity Manual* February (1998), pp. 1–56.

[55] D. V. d. M. Ruschmann, L. Paolucci, and N. A. L. Maciel. "Capacidade de carga no planejamento turístico: Estudo de caso da Praia Brava – Itajaí frente à implantação do Complexo Turístico Habitacional Canto da Brava." In: *Revista Brasileira de Pesquisa em Turismo* 2.2 (2008), pp. 41–63. DOI: 10.7784/rbtur.v2i2.102.

[56] P. N. Sadikin, H. S. Arifin, B. Pramudya, and S. R. Mulatsih. "Carrying capacity to preserve biodiversity on ecotourism in Mount Rinjani National Park, Indonesia." In: *Biodiversitas* 18.3 (2017), pp. 978–989. DOI: 10.13057/biodiv/d180316.

[57] M. Salemi, S. A. Jozi, S. Malmasi, and S. Rezaian. "Conceptual framework for evaluation of ecotourism carrying capacity for sustainable development of Karkheh protected area, Iran." In: *International Journal of Sustainable Development and World Ecology* 26.4 (2019), pp. 354–366. DOI: 10.1080/13504509.2019.1570379.

[58] I. Seidl and C. A. Tisdell. "Carrying capacity reconsidered: From Malthus' population theory to cultural carrying capacity." In: *Ecological Economics* 31.3 (1999), pp. 395–408. DOI: 10.1016/S0921-8009(99)00063-4.

[59] R. Silva. "A Tourism Overcrowding Sensor Using Multiple Radio Techniques Detection." Master's thesis. Iscte – Instituto Universitário de Lisboa, 2019.

[60] J. Sorgalla. "AjiL: A Graphical Modeling Language for the Development of Microservice Architectures." In: *Extended Abstracts of the Microservices 2017 Conference*. 2017.

[61] R. Streeton, M. Cooke, and J. Campbell. "Researching the researchers: using a snowballing technique." In: *Nurse researcher* 12.1 (2004), pp. 35–46. DOI: 10.7748/nr2004.07.12.1.35.c5929.

[62] I. W. Suana, H. Ahyadi, G. Hadiprayitno, S. Amin, L. A. T. T. W. S. Kalih, and F. X. Sudaryanto. "Environment carrying capacity and willingness to pay for bird-watching ecotourism in Kerandangan Natural Park, Lombok, Indonesia." In: *Biodiversitas* 21.5 (2020), pp. 2266–2274. DOI: 10.13057/biodiv/d210557.

[63] P. Taillandier, B. Gaudou, A. Grignard, Q. N. Huynh, N. Marilleau, P. Caillou, D. Philippon, and A. Drogoul. "Building, composing and experimenting complex spatial models with the GAMA platform." In: *GeoInformatica* 23.2 (2019), pp. 299–322. DOI: 10.1007/s10707-018-00339-6.

[64] M. Thyne, K. M. Woosnam, L. Watkins, and M. A. Ribeiro. "Social distance between residents and tourists explained by residents' attitudes concerning tourism." In: *Journal of Travel Research* (2020), p. 0047287520971052.

[65] A. Tordeux, M. Chraibi, and A. Seyfried. "Collision-free speed model for pedestrian dynamics." In: *Traffic and Granular Flow'15*. Springer, 2016, pp. 225–232.

[66] A. C. Tricco, J. Antony, W. Zarin, L. Strifler, M. Ghassemi, J. Ivory, L. Perrier, B. Hutton, D. Moher, and S. E. Straus. "A scoping review of rapid review methods." In: *BMC medicine* 13.1 (2015), p. 224. DOI: https://doi.org/10.1186/s12916-015-0465-6.

[67] G. Vizzari, L. Crociani, and S. Bandini. "An agent-based model for plausible wayfinding in pedestrian simulation." In: *Engineering Applications of Artificial Intelligence* 87 (2020), p. 103241.

[68] A. K. Wagoum, A. Tordeux, and W. Liao. “Understanding human queuing behaviour at exits: an empirical study.” In: *Royal Society open science* 4.1 (2017), p. 160896.

[69] World Tourism Organization (UNTWO), Centre of Expertise Leisure Tourism & Hospitality Sciences, NHTV Breda University of Applied Sciences, and NHL Stenden University of Applied Sciences. *‘Overtourism’? – Understanding and Managing Urban Tourism Growth beyond Perceptions, Executive Summary*. World Tourism Organization (UNWTO), 2018. DOI: 10.18111/9789284420070.

[70] V. Yepes Piqueras. “Ordenación y gestión del territorio turístico: las playas.” In: *Ordenación y gestión del territorio turístico.* January 2002. Tirant lo Blanch. 2002, pp. 549–579.

[71] D. A. Zacarias, A. T. Williams, and A. Newton. “Recreation carrying capacity estimations to support beach management at Praia de Faro, Portugal.” In: *Applied Geography* 31.3 (2011), pp. 1075–1081. DOI: 10.1016/j.apgeog.2011.01.020.

[72] S. Zhang, F. Zhen, B. Wang, Z. Li, and X. Qin. “Coupling Social Media and Agent-Based Modelling: A Novel Approach for Supporting Smart Tourism Planning.” In: *Journal of Urban Technology* (2020), pp. 1–19.

[73] Y. Zhuang, Z. Liu, A. Schadschneider, L. Yang, and J. Huang. “Exploring the behavior of self-organized queuing for pedestrian flow through a non-service bottleneck.” In: *Physica A: Statistical Mechanics and its Applications* 562 (2021), p. 125186.

Urban Tourism Crowding Dynamics:
Carrying Capacity and Digital Twinning

Duarte Almeida