Improving Emergency Medical Services Through Vehicle Location Optimization

Guilherme Costa Antunes Ferreira dos Santos

Dissertation to obtain the Master of Science Degree in

Industrial Engineering and Management

Supervisors: Prof. Inês Marques Proença
Prof. Ana Paula Ferreira Dias Barbosa Póvoa

Examination Committee

Chairperson: Prof. Carlos António Bana e Costa
Supervisor: Prof. Inês Marques Proença
Member of the Committee: Prof. Miguel Jorge Vieira

December 2019

“All models are wrong; some models are useful.”

\- *George E.P. Box*

Declaration

I declare that this document is an original work of my own authorship and that it fulfils all the requirements of the Code of Conduct and Good Practices of the Universidade de Lisboa.

Declaração

Declaro que o presente documento é um trabalho original da minha autoria e que cumpre todos os requisitos do Código de Conduta e Boas Práticas da Universidade de Lisboa.

ABSTRACT

Emergency Medical Services (EMS) are paramount in saving lives and their importance has long been recognized. The Portuguese EMS, SIEM, has evolved since the 1960's into a complex system, comprising several stages and entities. The National Institute for Medical Emergencies (INEM) is responsible for managing the SIEM. Among other decisions, it is responsible for locating emergency vehicles, which directly impacts the medical outcomes of the population. Nonetheless, current vehicle locations are a result of incremental modifications guided by experience. In this context, more sophisticated decision-support methods could be beneficial.

A Multi-Objective Dynamic Mixed-Integer Programming model is developed to support decisions concerning the station selection, the allocation of vehicles to stations and the assignment of demand to vehicles. In general, the model considers multiple vehicles and call priorities, as well as micro and macro time periods, to determine how the existing system can be gradually improved. Three objectives are considered: coverage, cost and equity. Besides, a hybrid heuristic based on problem decomposition is proposed to streamline model solution for the first objective.

Application of the model to two regions suggests that, under the current restrictions, only slight improvements are possible. The model is further explored to study fleet expansion and seasonal vehicle allocation decisions, aligned with current planning practices. Alternative operating scenarios show that, with the current resources, improvements of up to 13.3% in coverage could be attained by relocating stations and vehicles simultaneously. Finally, computational tests of the heuristic prove it to be an effective solution procedure for large instances.

Keywords: Emergency Medical Services; Ambulance Location; Optimization; Hybrid Heuristic; Mixed-Integer Programming.

RESUMO

Os serviços de emergência médica (EMS) são fulcrais para garantir o bem-estar da população e a sua importância é reconhecida há vários anos. O Sistema Integrado de Emergência Médica (SIEM), estabelecido na década de 1960, evoluiu para um sistema complexo, operando múltiplos veículos e diferentes estágios de intervenção. O Instituto Nacional de Emergência Médica (INEM) é responsável por gerir o SIEM. Entre outras decisões, compete-lhe posicionar os meios de emergência, com um impacto direto na saúde da população. Ainda assim, a configuração atual do sistema resulta de decisões incrementais nunca sujeitas a uma avaliação holística. Neste contexto, o INEM beneficiaria de métodos científicos que permitissem informar os processos de decisão.

Por este motivo, um modelo Dinâmico Multiobjectivo de Programação Inteira Mista é desenvolvido para apoiar decisões relativas à seleção de estações, à afetação de veículos e à satisfação da procura. Em geral, o modelo considera múltiplos veículos e prioridades de emergência, bem como diversos períodos temporais e três objetivos: cobertura, custo e equidade. Uma heurística híbrida é proposta para acelerar a resolução do modelo para o primeiro objetivo.

Aplicação do modelo a duas regiões sugere que, cumprindo as restrições atuais, pequenas melhorias são possíveis. O modelo também é utilizado para testar decisões sobre a expansão da frota e veículos sazonais. Cenários alternativos sugerem que, com os recursos atuais, melhorias de cobertura até 13.3% podem ser obtidos. Finalmente, resultados computacionais da heurística mostram que é um método de solução eficaz para instâncias grandes.

Palavras-chave: Serviços de Emergência Médica; Localização de Ambulâncias; Otimização; Heurística Híbrida; Programação Matemática

AKNOWLEDGEMENTS

At the end of this long journey, I take the opportunity to deeply thank everyone who has made it possible for me to reach this stage successfully.

Firstly, I am very thankful to Professor Inês Marques and Professor Ana Póvoa for their constant support and guidance. Their invaluable insights challenged me to go further at each step, but it is their constant availability and the many wise advices which I will take with me for the rest of my life.

I would also like to thank INEM for the opportunity to develop this project in a real-life setting. In particular, I thank Dr. João Reis for his availability and eagerness in discussing the challenges and opportunities of the problem, and his determination in seeing the project successfully concluded.

I also wish to express my deepest gratitude to all my friends and colleagues at IST. You have made this incredible journey much easier and happier, and your good mood and sincere friendship are the best things I take from my time at IST. I also want to thank all my friends, especially those from MarCha, for growing with me and being my second family.

Last, but surely not the least. To my parents, for their unconditional support and love throughout my life and for teaching me what really matters. To my brother Duarte, for being my role-model since I can remember and showing me how to be a great human-being. To Marta, for her constant support, understanding and unshakeable belief in me throughout all these years. Without you, it would not have been possible.

\---

E, ainda, uma palavra muito especial para a minha maior referência, a minha Avó Rita, e para o meu grande amigo, Pedro Paulino. Um obrigado não chega... Espero que estejam orgulhosos.

\---

This work was developed within the scope of project DSAIPA/AI/0044/2018 funded by the Portuguese National Science Foundation (FCT).

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACCRONYMS

AEM	*Ambulância de Emergência Médica* (Medical Emergency Ambulance)
AHQM	Approximate Hypercube Queueing Model
ALS	Advanced Life Support
BACOP	Back-up Coverage Problem
BB	Branch-and-Bound
BIP	Binary Integer Programming
BLS	Basic Life Support
CODU	*Centro de Orientação de Doentes Urgentes* (Urgent Patient Dispatching Centre)
DSM	Double Standard Model
EMS	Emergency Medical Services
FAST	Fire and Ambulance Service Technique
FLEET	Facility-Location Equipment-Emplacement Technique
GA	Genetic Algorithm
GPCG	*Gabinete de Planeamento e Controlo de Gestão* (Planning and Management Control Office)
HQM	Hypercube Queueing Model
ILS	Immediate Life Support
INEM	*Instituto Nacional de Emergência Médica* (National Institute for Medical Emergency)
IP	Integer Programming
LR	Linear Relaxation
LSCP	Location Set Covering Problem
MALP	Maximum Availability Location Problem
MCLP	Maximum Covering Location Problem
MEM	*Motociclo de Emergência Médica* (Medical Emergency Motorcycle)
MEXCLP	Maximum Expected Coverage Location Problem
MIP	Mixed-Integer Programming
MOFLEET	Multi-Objective Facility-Location Equipment-Emplacement Technique
NHS	National Health Service
NINEM	*Ambulância Não-INEM* (Non-INEM Ambulance)
NLP	Non-Linear Programming
OR	Operations Research
PEM	*Ambulâncias de Postos de Emergência Médica* (Medical Emergency Station Ambulances)
PLSCP	Probabilistic Location Set Covering Problem
PROFLEET	Probabilistic Facility-Location Equipment-Emplacement Technique

PSP	*Polícia de Segurança Pública* (Public Safety Police)
Q-PLSCP	Queueing Probabilistic Location Set Covering Problem
RES	*Ambulâncias de Postos de Reserva* (Reserve Station Ambulances)
SA	Simulated Annealing
SHEM	*Serviço de Helicópteros de Emergência Médica* (Medical Emergency Helicopter System)
SIEM	*Sistema Integrado de Emergência Médica* (Integrated Medical Emergency System)
SIV	*Ambulância de Suporte Imediato de Vida* (Immediate Life Support Ambulance)
SUB	*Serviço de Urgência Básica* (Basic Urgency Service)
SUMC	*Serviço de Urgência Médico-Cirúrgica* (Medical-Surgical Urgency Service)
SUP	*Serviço de Urgência Polivalente* (Multipurpose Urgency Service)
TEAM	Tandem Equipment Allocation Model
TEPH	*Técnico de Emergência Pré-Hospitalar* (Pre-Hospital Emergency Technicians)
TETRICOSY	Telephonic Triage and Counseling System
TIMEXCLP	Time-Dependent Maximum Expected Coverage Location Problem
TIP	*Ambulâncias de Transporte Inter-Hospital Pediátrico* (Inter-Hospital Pediatric Transport Ambulances)
UBUL	Upper-Bound Unavailability Location Model
UMIPE	*Unidade Móvel de Intervenção Psicológica de Emergência* (Mobile Unit of Phycological Emergency Intervention)
VNS	Variable Neighborhood Search

1. INTRODUCTION

The aim of this chapter is to introduce this dissertation, motivating the problem to be addressed and the goals of this research. Additionally, the research methodology and the structure of the dissertation are outlined. This chapter is divided into three sections: section 1.1 provides context regarding the problem under study, section 1.2 outlines the main goals of the dissertation and section 1.3 presents the structure of the remaining chapters.

1.1 PROBLEM BACKGROUND AND MOTIVATION

Emergency Medical Services (EMS) systems are designed to save lives and they are a vital component of pre-hospital medical care (World Health Organization, 2005). Since the 1960's, they have been recognized as an important system with a direct impact on the medical outcomes of emergency patients (Krafft *et al.*, 2003). Therefore, they have a significant influence on the overall welfare of the population.

These systems are usually complex, operating multiple tiers of vehicles, which interact with dispatching centres and hospitals (Jagtenberg, 2016). Even though there are different EMS typologies around the world, a common objective for EMS planners is to provide quality health care while keeping costs low (Reuter-Oppermann, Van Den Berg and Vile, 2017). In particular, ensuring short response times has long been recognized as a critical factor guiding decision-making (Li *et al.*, 2011). In order to provide quality care, planners must make multiple planning decisions regarding resources, such as staff, the structure of the emergency fleet and emergency station locations (Bélanger, Ruiz and Soriano, 2018).

The Portuguese integrated EMS system, *Sistema Integrado de Emergência Médica* (SIEM), is no exception. The SIEM is managed by the National Institute for Medical Emergency, *Instituto Nacional de Emergência Médica* (INEM). Its main goal is to provide prompt and appropriate medical care to victims and patients in continental Portugal. The SIEM has evolved since the 1960's, and nowadays comprises multiple entities with varying levels of responsibility, several types of emergency vehicles and emergency priorities, answering more than 1.300.000 calls per year, around 90% of which require direct medical intervention. Nevertheless, despite having to manage such a challenging system, INEM planners must often rely on experience and intuition to make difficult planning decisions under uncertainty, limited by budget constraints and balancing multiple stakeholders and objectives.

One of these decisions concerns the selection of long-term positions of emergency vehicles while they are idle. This decision actually comprises two related issues: selecting bases to be used and allocating existing vehicles to those bases. Current vehicle positions have been progressively developed as the system was expanded and new vehicles acquired. These decisions have been mostly supported by intuition and experience. Therefore, there is no guarantee that the implemented solution makes the best use of available resources to meet the goals of the SIEM and maximize population welfare. Furthermore, changing the existing configuration of vehicles can have costs, both financial and in human lives, meaning that INEM planners are naturally reluctant to make experiments without a solid reason to believe that these will be effective.

In this context, the use of more sophisticated methods to support decision-making could potentially lead to more effective and efficient solutions and contribute to enhance the overall performance of the SIEM.

In fact, the limitations of the current planning methods and need for a scientific model-based approach have been recognized in 2010 by the Portuguese Audit Office (Tribunal de Contas, 2010).

1.2 DISSERTATION GOALS

This setting motivates the present study, whose ultimate goal is to develop and apply Operations Research (OR) techniques to improve INEM's emergency vehicle location planning processes. This project is part of a series of collaborative projects between INEM and the Centre of Management Studies (CEG) of Instituto Superior Técnico which focus on different EMS planning problems.

Ultimately, the objective is to develop an optimization model that is able to produce recommendations regarding vehicle location at INEM. This model should be flexible, integrating multiple objectives while providing a realistic representation of the SIEM. Furthermore, and despite being inspired by INEM's case study, the proposed model should be general enough to allow modelling different EMS systems within a single framework, providing insights into the impact of alternative policies in the system's performance. Secondary research goals include:

- Characterizing EMS systems, particularly the SIEM, focusing on the most important components of EMS operation, highlighting their complex nature;
- Surveying current planning practices used by EMS practitioners;
- Reviewing previous research on the field, centred around specific modelling approaches and their consequences, leveraging these studies to support the development of the model;
- Identifying potential gaps in the literature related to the needs of the case-study and proposing alternative modelling approaches;
- Contributing to the existing literature by modelling, formulating and implementing an optimization model in a real EMS context;
- Producing recommendations to support INEM in their vehicle location planning decisions.

1.3 RESEARCH METHODOLOGY

To achieve the above, the proposed research methodology includes five fundamental steps, as outlined in Figure 1.

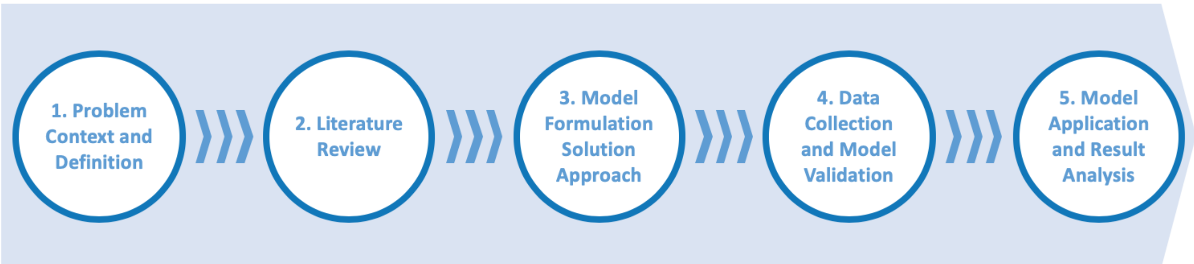


Figure 1 - Proposed methodology.

- **Step 1 – Case-Study Description and Problem Definition**

  This step seeks to introduce the EMS system which serves as a case-study for this dissertation: the SIEM. By highlighting its main objectives and features, and describing existing planning processes and performance indicators, an overview of the main challenges of vehicle planning at INEM is provided and the problem statement can be refined.

- **Step 2 – Literature Review**

  The goal of the literature review is to survey the state-of-the-art of Location Science concerning emergency vehicle positioning. A brief overview of the typical response process, associated metrics and planning problems provides the theoretical foundation to present the wide range of optimization models available in the literature, focusing on coverage, discrete models for strategic planning, while also mentioning relevant solution techniques and applications.

- **Step 3 – Model Formulation and Solution Approach**

  Considering the problem definition and relevant models in the literature, assumptions are stated, and a model is formulated. The Mathematical Programming model is developed under collaboration with INEM, to ensure that it adequately addresses the needs of EMS systems. For this purpose, interviews are required to capture different aspects of the decision-making model. Preliminary tests are conducted using a general-purpose solver. Furthermore, a solution approach exploring model structure is developed to streamline model solution.

- **Step 4 – Data Collection and Model Validation**

  Relevant data for the model is collected with INEM and treated. The model is then tested on real data to ensure its validity. Here, the impact of each major assumption should be tested. Again, this step requires strict collaboration with INEM and, likely, iteration with the previous step to refine the model until it accurately describes EMS systems. This step is paramount to promote acceptance of the model by INEM.

- **Step 5 – Model Application and Result Analysis**

  The model is applied to the case-study of SIEM to produce recommendations. Results are then analysed and discussed, and experiments are made to understand the impact of different constraints and scenarios on the performance of the system.

1.4 STRUCTURE OF THE DISSERTATION

This dissertation is structured in 8 chapters:

- **Chapter 1 – Introduction**

  Corresponding to the present chapter, it introduces and motivates the topic of this study, highlights the main goals to be achieved and outlines the proposed research methodology and structure of the document.

- **Chapter 2 – What is an EMS system?**

  Provides an overview of EMS systems, their purpose and structure. Key concepts are introduced, which help grasp the complex nature of these systems. A typical response process is detailed, and planning decisions are outlined at all levels of decision-making: strategic, tactical and operational.

- **Chapter 3 – The Portuguese EMS system: SIEM**

  The third chapter seeks to describe the case-study addressed in this dissertation. INEM is introduced and the SIEM is described, including existing emergency vehicles. The planning process for emergency vehicle location is also presented.

- **Chapter 4 –Literature Review**

  The objective of chapter 4 is to survey the existing literature on OR methods, specifically optimization models, which can be applied to support EMS vehicle location planning. Several types of models are reviewed, focusing on how specific components of operation are modelled.

- **Chapter 5 – Model Formulation**

  In chapter 5, insights gathered from the literature review and the description of the case study are leveraged to refine the problem statement and develop an optimization model aimed at supporting the emergency vehicle location planning process at INEM. The main modelling approaches and assumptions are detailed. Additionally, a solution process is proposed to enable the application of the model to large instances considering the primary model objective.

- **Chapter 6 – Data Collection and Treatment**

  This chapter details the data collection and treatment procedures required to apply the model to the case study. Underlying limitations and necessary assumptions in the data are described. Summary tables are presented with the input parameters resulting from these procedures.

- **Chapter 7 – Case-Study Results**

  Describes the implementation and computational experiments performed with the proposed model and solution approach. The main conclusions and findings of these experiments are described and recommendations for INEM are presented.

- **Chapter 8 – Final Conclusions and Future Work**

  The last chapter summarises the most relevant features and conclusions of this study, highlighting opportunities for future research.

2. WHAT IS AN EMS SYSTEM?

The aim of this chapter is to introduce EMS systems by presenting the main features of a standard EMS system and introduce terminology and concepts which are useful for the remainder chapters. Section 2.1 defines the purpose of EMS systems and describes two basic categories, section 2.2 typifies a typical response process and relevant performance metrics, section 2.3 reviews planning decisions for EMS planners and, finally, section 2.4 presents the chapter's conclusions.

2.1 PURPOSE AND STRUCTURE OF EMS SYSTEMS

Prompt and effective pre-hospital medical care is essential in saving lives. Therefore, most countries have established EMS over the years (World Health Organization, 2005). EMS are provided via EMS systems. Starting from simple structures dedicated to military conflicts, they have developed into complex systems with interacting vehicles (ambulances, motorcycles, helicopters, cars, etc.), dispatching centres and health care facilities (hospitals, health centres) (Jagtenberg, 2016).

All EMS systems exist to save lives by providing quality medical assistance to injured people in the shortest possible time. However, each country organises EMS activities differently, and there is no common standard for EMS systems (Fischer *et al.*, 2011; Totten and Bellou, 2013). This diversity results from the fact that EMS systems were only recognized as such – that is, systems – very recently (Krafft *et al.*, 2003).

Nevertheless, two main EMS system categories can be identified (Dick, 2003):

- **Anglo-American System:** the patient is brought to the doctor. The goal is to respond to calls quickly: trained paramedics provide minimal intervention at the scene and transport victims directly to an appropriate medical facility;
- **Franco-German System:** the doctor is brought to the patient. Specialized paramedics and emergency physicians are deployed so that life-threatening emergencies are treated on scene and during transportation. Complex treatment can be administered, including defibrillation, intubation and life-saving drugs. In this model, transportation speed is less crucial.

However, nowadays most EMS systems do not fall into a single category, instead mixing features of both philosophies (Al-Shaqsi, 2010).

2.2 A TYPICAL EMS RESPONSE PROCESS AND PERFORMANCE METRICS

Although systems vary significantly, a "typical EMS system", as described by Reuter-Oppermann, Van den Berg and Vile (2017), employs different types of vehicles to respond to emergencies with different priorities. The typical response process is shown in Figure 2.

The system is activated through a call to an emergency number (e.g. 112 in Europe, 911 in the USA) and is answered by a dispatching centre. Here, several models exist as to how the call is redirected to the competent service (EENA, 2018). A triage process follows, in which the location and severity of the emergency is assessed. Depending on this stage, the dispatcher decides whether it is necessary to deploy emergency resources. If emergency vehicles are required, they are dispatched from their stations following some dispatching rule (usually, the closest idle vehicle). The vehicle travels to the scene, provides varying degrees of emergency care and, if necessary, transports the patient to an appropriate health unit. After transferring the patient, and potentially restocking medical supplies, the vehicle becomes idle and returns to its base or, if relocation is allowed, to another base. In some cases, if no vehicle is available, a call can be queued – this is, however, very rare (Jagtenberg, 2016).

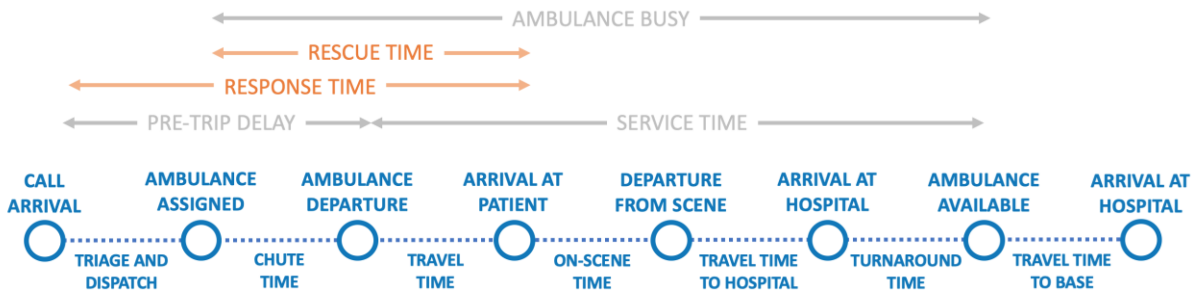


Figure 2 - Typical response process of an EMS system. Source: Adapted from Reuter-Oppermann, Van den Berg and Vile (2017)

Emergency vehicles can be staffed and equipped differently, thus providing variable care levels. There is considerable debate regarding the most effective vehicle mix of EMS systems, with some authors supporting tiered systems (i.e. with multiple vehicles), while others preferring single vehicle systems. Either way, it is typical to consider two vehicle categories, according to the care level provided (Chong, Henderson and Lewis, 2016):

- **Advanced Life-Support (ALS) vehicles**: operated by specialized staffed (physician, nurses, paramedics) and equipped with advanced medical devices, they are more expensive but provide a higher level of care;
- **Basic Life-Support (BLS) vehicles**: staffed by technicians (e.g. Emergency Medical Technicians), providing a lower range of treatments at a lower cost.

Measuring the performance of an EMS system is also an important topic. Typically, EMS systems are evaluated on time interval metrics, being Response Time the most popular (Reuter-Oppermann, Van Den Berg and Vile, 2017). Some of these metrics are represented in Figure 2. In many cases, EMS legislation sets Response Time targets: in the UK, life-threatening emergencies should be answered in 8 minutes (NHS England, 2017), while the Response Time limit in the Netherlands is 15 minutes for all emergencies (Wulterkens, 2005). Portuguese legislation does not establish such standards. Still, INEM uses a less demanding time interval metric, called Rescue Time, as a performance indicator.

These metrics are attractive because they are simple to understand, easily measurable, comparable and objective, but they may lead to unintended consequences and be inadequate to modern reality. Another problem is the lack of commonly accepted descriptions for each of these time intervals, leading to different EMS systems using the same name for different metrics (Myers *et al.*, 2008).

Nevertheless, additional performance measures have been proposed. For instance, outcome-based measures, depending on the patient's clinical situation, have been developed (Myers *et al.*, 2008). Furthermore, the European Emergency Data Project tried to identify common indicators and establish an evaluation framework of European EMS systems. The resulting metrics include: Available unit hours per 100.000 habitants; Response Time (% under 8 minutes) of critical calls; Rate of high priority responses per 100.000 habitants; Rate of First Hour Quintet[1] incidents per 100.000 habitants; ALS interventions per 100.000 habitants per year (Fischer *et al.*, 2011).

2.3 PLANNING ISSUES FOR EMS SYSTEMS

To run effective EMS systems, planners must make a series of complex decisions, involving multiple resources, such as emergency vehicles, emergency stations, crews and dispatchers. They usually seek to provide quality health care at a bearable cost, leading to a natural trade-off (Jagtenberg, 2016). These planning decisions are traditionally classified into three levels (Reuter-Oppermann, Van Den Berg and Vile, 2017; Bélanger, Ruiz and Soriano, 2018):

- Strategic decisions: long-term, holding for at least one year, but eventually spanning for several decades;
- Tactical decisions: medium-term, valid for periods ranging from one month to one year;
- Operational decisions: made for a daily basis, or even in real time.

Figure 3 summarizes some of the decisions that EMS planner face at each decision level.

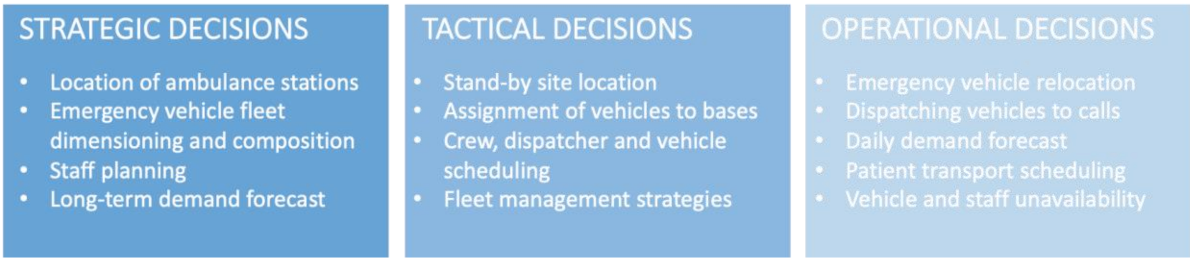


Figure 3 - Planning levels and examples of problems for EMS systems. Adapted from: Reuter-Oppermann, van den Berg and Vile (2017) and Bélanger, Ruiz and Soriano (2018).

EMS planners face challenges regarding vehicle location across all levels of decision-making: the size of the emergency fleet, its structure and which emergency stations to open (strategic); assignment of vehicles to stations (tactical); vehicle dispatching and relocation (operational) (Bélanger, Ruiz and Soriano, 2018). As discussed in the next chapter, this dissertation concerns strategic and tactical decisions. Besides, vehicle-related decisions often require other important tasks, such as demand forecasting or travel time analysis (Reuter-Oppermann, Van Den Berg and Vile, 2017). Fortunately, the combined use of modern information systems, mathematical optimization models, solution methods and software can help make better decisions, supported by *a priori* testing of different solutions (Beraldi, Bruni and Conforti, 2004).

2.4 CHAPTER CONCLUSIONS

EMS systems are very important components of emergency care, contributing to the population's welfare. They are usually complex, with significant variations across countries, meaning that EMS planners face difficult decisions regarding resource management which must account for the particularities of each region. Now that general concepts of EMS systems have been defined, the case-study addressed by this dissertation is described in Chapter 3. Existing literature on strategic and tactical emergency vehicle location planning in then reviewed Chapter 4.

[1] First Hour Quintet incidents include cardiac arrest, severe respiratory failure, severe trauma, stroke and chest pain.

3. THE PORTUGUESE EMS SYSTEM: SIEM

The aim of this chapter is to describe the case study addressed in this dissertation. Section 3.1. describes INEM, the entity responsible for the Portuguese EMS. Section 3.2. describes how the Portuguese EMS system, SIEM, operates and presents the existing types of vehicles. Section 3.3. discusses vehicle location planning at INEM, leading to a refined statement of the research problem in section 3.4. Finally, section 3.5 presents the chapter's conclusions.

3.1 THE NATIONAL INSTITUTE FOR MEDICAL EMERGENCY (INEM)

3.1.1 MISSION AND VISION

INEM is the public entity under the Ministry of Health responsible for managing the Portuguese EMS. INEM's mission was lastly updated in Decree-Law 34 of 2012, which formally defined it as "*guaranteeing the operation, in continental Portugal, of an Integrated Emergency Medical System (SIEM), ensuring prompt and correct provision of medical care to victims and patients (…)*". To fulfil its purpose, INEM must define, organize, participate in, coordinate and evaluate SIEM's activities and guarantee their articulation. INEM's responsibilities include (Instituto Nacional de Emergência Médica, 2017a):

- Call reception, triage, counselling and vehicle dispatching;
- Providing pre-hospital emergency care (medicalized and non-medicalized) and coordinating all entities to offer an integrated response (e.g. hospitals, police officers, firefighters, …);
- Patient transportation, referral, reception and treatment at the hospital;
- Defining, planning, providing and certifying medical emergency training activities;
- Civil planning, prevention and raising awareness towards medical emergency;
- Establishing and maintaining a medical telecommunication network.

INEM's vision is *"to be innovative, sustainable, motivating and the reference organization in emergency medical care"*. The organization seeks to develop new processes, technologies and skills, while constantly monitoring its performance and continuously improving its operations to deliver better aid (Instituto Nacional de Emergência Médica, 2017a).

3.1.2 ORGANIZATIONAL STRUCTURE

INEM is organized around three decentralized Regional Delegations (North, Centre, South) which articulate with centralized Organic Units, as presented in Figure 4. Both centralized and decentralized units report to the Executive Board. Decentralized units are responsible for operations management and coordination within their geographical area. Centralized units provide services for the whole territory on three areas of activity: operational, logistics support and management support.

Vehicle planning, which is the focus of this study, is under the supervision of the Planning and Management Control Office (*Gabinete de Planeamento e Controlo de Gestão,* GPCG). This office produces studies and recommendations regarding vehicle location and delivers them to the board. The final decision is always left to the board, which often needs to balance political goals as well as the recommendations produced by the GPCG.

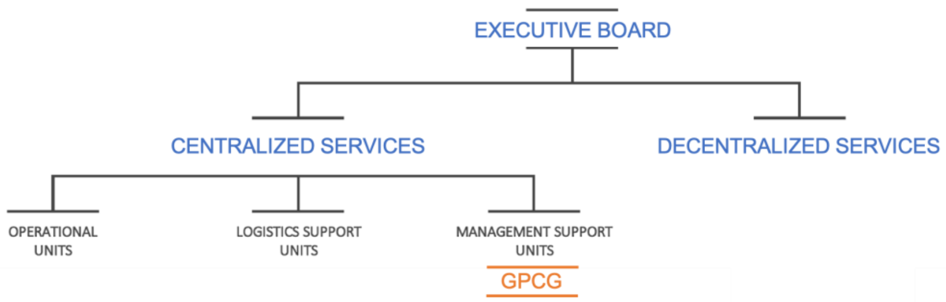


Figure 4 - INEM's organisational structure.

3.1.3 MILESTONES OF PORTUGAL'S EMS SYSTEM

Portugal's EMS system dates back to 1965, when the first emergency number – 115 – was established in Lisbon. In this system, an ambulance, operated by two police officers, transported the victims of emergencies to the nearest hospital. In subsequent years, this service was extended to other regions of Portugal, such as Faro, Oporto and Coimbra (Instituto Nacional de Emergência Médica, 2013).

Around 1980, Portugal started building its integrated EMS system, called SIEM. INEM was later established, in 1981, as the public body "*responsible for coordinating the activities of Medical Emergency to be executed by the several entities*" (Ministérios da Defesa Nacional, das Finanças e do Plano, dos Assuntos Sociais e da Reforma Administrativa, 1981). The existing National Ambulance System was integrated into INEM. In the following years, the SIEM expanded rapidly and deals were signed with public entities, including firefighters, police and the Red Cross (Gomes *et al.*, 2004). The goal was to define responsibilities and promote coordination, thus achieving an integrated response. In 1987, the first Dispatching Centre, *Centro de Orientação de Doentes Urgentes* (CODU), was created.

From 1990 to 2006, four new CODUs were opened, becoming responsible for all medical calls from the European Emergency Number – 112. New types of vehicles were acquired: Helicopters, VIC (*Viatura de Intervenção em Catástrofe*) and VMER (*Viatura Médica de Emergência e Reanimação*). During this period, INEM started training emergency technicians required by the SIEM (Instituto Nacional de Emergência Médica, 2013). Between 2007 and 2017, Immediate Life Support (ILS) ambulances (SIV) were implemented and additional BLS vehicles were acquired. By 2017, INEM had accomplished its goal of locating an ambulance in every municipality of Portugal. Several training initiatives were carried out, and an effort was made to modernize INEM's information systems. During this period, INEM's processes and responsibilities (as well as its partners) have been progressively redefined (INEM, 2017).

3.2 THE INTEGRATED EMERGENCY MEDICAL SYSTEM: SIEM

The SIEM comprises several coordinated activities performed by different bodies within the scope of medical emergency. These activities include pre-hospital medical care, transportation, hospital reception, patient referral and treatment at receiving health unit. The main entities involved in delivering emergency medical care include INEM, firefighters, police officers, the Red Cross, doctors, nurses and hospitals. Consequently, coordination is crucial to ensure an effective and resource-efficient intervention, which is the main goal of the SIEM.

Firstly, an overview of SIEM's philosophy and stages is provided, before describing the most crucial stages, call reception and triage, emergency vehicles, dispatching rules and system coordination.

3.2.1 OVERVIEW: SIEM'S PHILOSOPHY AND STAGES

The SIEM follows the Franco-German model. It works as an extension of emergency departments of National Health Service (NHS) hospitals. The goal is to perform as many diagnostic and treatment procedures as possible in the shortest time. This enables the victim's stabilization and early treatment. Pre-hospital diagnosis is important since it contributes to select the most appropriate health unit to receive the patient (Instituto Nacional de Emergência Médica, 2013, 2015).

There are several stages of SIEM's intervention, similar to the ones identified in the previous chapter. Firstly, the emergency situation is detected, usually by civilians, and reported by calling 112. After screening and evaluating the call's priority, the appropriate vehicle is dispatched. While the emergency vehicle is travelling to the scene, basic first-aid care can be applied by the caller. Once the vehicle arrives, medical care is provided to stabilize the victim and start treatment. If necessary, the patient is transported to an appropriate health unit, while care continues in-transit. Finally, the patient is transferred to the receiving health unit so that treatment can be finalized. The vehicle and its crew are then released. These stages are represented by the INEM's Star of Life symbol (Instituto Nacional de Emergência Médica, 2013). Alongside INEM, firefighters are major EMS providers, especially in rural areas. The Red Cross is also an important provider, complementing both INEM and the firefighters.

Non-urgent patient transportation is not part of INEM's responsibilities and, hence, is not addressed in this study. This service is provided by firefighters, the Red Cross, municipalities or private companies (IPSS), dully authorized by INEM (República Portuguesa, 2014; INEM, 2015).

3.2.2 CALL RECEPTION AND TRIAGE

112-calls are answered in two Operational Centres managed by *Polícia de Segurança Pública* (PSP). PSP operators evaluate the call and forward it to the competent authority. If they face a health-related emergency, the call is redirected to one of INEM's CODU. INEM operates four CODUs – Porto (North), Coimbra (Centre), Lisbon (South) and Faro – but only the first three answer calls. They ensure 24/7 coverage of emergency calls in Portugal. Their job is to establish a connection between the emergency and SIEM, deploying the necessary resources and entities to ensure an adequate response (Instituto Nacional de Emergência Médica, 2017b). Besides 112 calls, the CODUs also answer two other requests: calls transferred from *Linha Saúde 24* (24 Health Line, which provides medical advice to the population through the telephone) and inter-hospital emergency transportation requests, both of which are treated as 112 calls. In 2017, the CODUs answered about 1 368 141 calls, around 3 748 per day.

Emergency calls are answered by technicians trained by INEM, *Técnicos de Emergência Pré-Hospitalar* (TEPH), supported by a team of doctors and psychologists. A call is answered by the TEPH who has been idle for the longest, regardless of the caller's location or the TEPH's CODU. Although georeferencing information is already given by the 112 Operational Centers, the location of the emergency is always confirmed. The TEPH then proceeds to the triage stage, by asking questions to the caller, in order to assess the severity of the emergency and whether the victim's life is at risk. Since 2012, these questions are supported by a triage software developed by INEM, called TETRICOSY (Telephonic TRIage and COunseling SYstem) (Gerardo, 2017). As the TEPH records the caller's answers in the software, the priority level changes, and new questions are suggested. TETRICOSY employs different triage algorithms, depending on the emergency type, to suggest new questions. This procedure eases the TEPH's job, ensuring a standardized triage procedure. Based on the caller's responses, one of four priority levels can be assigned, as described in Figure 5.

Figure 5 - Priority levels description.

Figure 6 shows the evolution of monthly calls since 2013. It shows that SIEM's activity is driven primarily by P3 occurrences. In fact, the majority of medical emergencies - around 70-75% - are classified as P3, while only 10-15% are P1, 7-12% are P5 and the remaining have other priorities. It also exposes a trend of increasing P3 calls throughout the years and demand peaks during the summer and new year.

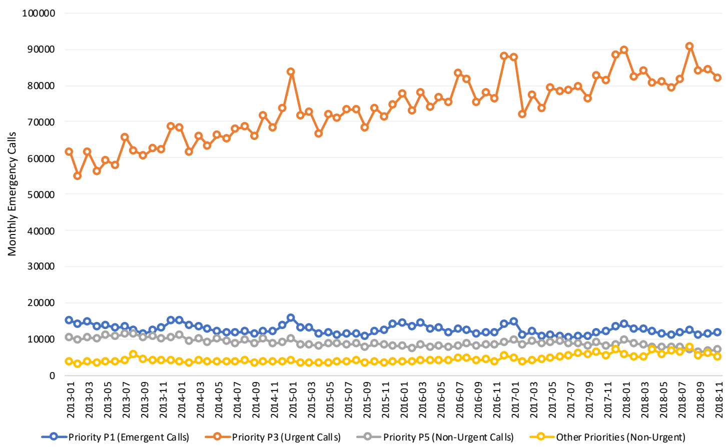


Figure 6 - Medical Emergencies Per Priority Level

After triage, if the situation is non-urgent, the call is transferred to *Linha Saúde 24*. If immediate assistance is required, the information collected during triage is made available to another team of TEPHs responsible for dispatching the appropriate emergency vehicles, following INEM's protocols.

3.2.3 EMERGENCY VEHICLES

The SIEM operates multiple types of emergency vehicles. Some are owned and operated by INEM, while others are financed by INEM but are the responsibility of a SIEM partner. Most vehicles work by shifts, since emergency calls are not evenly distributed through the day. All vehicles are described in Appendix A, including their name, description and number of vehicles from 2011 to 2018 (Instituto Nacional de Emergência Médica, 2017c). Since the location of emergency vehicles is the primary focus of this dissertation, this issue is discussed separately in section 3.3.

The basis of the system is a network of BLS ambulances, staffed by specialized technicians and capable of providing basic care and defibrillation. There are four types of BLS ambulances - AEM, PEM, RES and NINEM - complemented by BLS motorcycles (MEM). PEM and RES are operated by firefighters/Red Cross crews but financed by INEM, while NINEMs are owned and operated by a partner and AEMs are operated by INEM. BLS vehicles can respond to non-critical situations, but when the victim is emergent, differentiated care is required. For this purpose, two types of ALS vehicles exist: VMER and Helicopters. They have advanced equipment and are staffed by trained doctors and nurses. It is important to highlight that a VMER's crew is always part of a hospital's staff, which is responsible for crew scheduling and compensation. Another category of differentiated vehicle, the Immediate Life Support (SIV) ambulance, is staffed by a technician and a nurse, and can provide ILS. Additionally, there are other types of vehicles for special situations, such as psychological support (UMIPE) or situations involving children (TIP).

Although the overall size of the fleet has remained constant (from 644 vehicles in 2011 to 657 in 2018), its composition has changed. The number of PEMs has increased, while the number of RES and NINEM has decreased. This is due to the on-going conversion of RES into PEM and the establishment of new PEMs (Instituto Nacional de Emergência Médica, 2016b). This trend reflects INEM's effort to establish closer collaborations with its partners. INEM's vehicles have remained constant during this period.

It is also important to highlight the significative role played by vehicles owned and/or operated by firefighters and the Red Cross. In fact, PEM, RES and NINEM vehicles correspond to 75-77% of the emergency vehicle fleet, which is why coordinating with these stakeholders is fundamental for SIEM's effectiveness. Finally, INEM is currently under the process of renewing and expanding its fleet to improve coverage (Instituto Nacional de Emergência Médica, 2017b). This may be achieved by increasing the vehicle's availability but also through enhanced positioning.

3.2.4 VEHICLE SELECTION AND DISPATCHING

Depending on the priority, location and accessibility of the emergency, the TEPHs responsible for dispatching decide which vehicles are most appropriate. This decision is validated by a regulating doctor. Dispatching decisions can be made at any of the four CODUs.

Critical situations always require both a BLS vehicle and a differentiated vehicle, namely ILS/ALS vehicles, while less urgent situations require only BLS vehicles. Some emergency situations require pediatric care or a psychologist, in which case a TIP or an UMIPE is dispatched, respectively. Table 1 summarizes vehicle requirements for different priorities.

Table 1 - Vehicle types dispatched based on call priority.

Priority Level	**Vehicle Types**	**Possible Combinations of Vehicles**
P1 – Emergent Situations	BLS + ALS/ILS	AEM/PEM/RES/NINEM + VMER/SIV VMER + SIV SHEM*
P3 – Urgent Situations	BLS	AEM or PEM or RES or NINEM or SIV** MEM***
P5 – Non-urgent Situations	No vehicle	
Other Priorities (CIAV, P4 *autoridade*, P4 CDOS …)	UMIPE or other	Variable

*Only in exceptional situations.
**Although SIVs can be dispatched to P3 emergencies, this corresponds to only 4-6% of the total number of dispatches of SIVs (Instituto Nacional de Emergência Médica, 2018).
***MEMs are dispatched to P3 occurrences to quickly reach the occurrence and rule out false alarms. If the MEM determines that the call is valid, and transportation is required, another vehicle must be dispatched.

Figure 7 shows the number of dispatches for each vehicle type, per unit per year. It becomes clear that an AEM ambulance is dispatched more often than any other BLS ambulance (6-8 times per ambulance per day), followed by PEM, RES and NINEM ambulances, in line with the dispatching rule described previously. It is also clear that VMERs are more often dispatched than SIVs. UMIPEs, TIPs and SHEMs are seldom used.

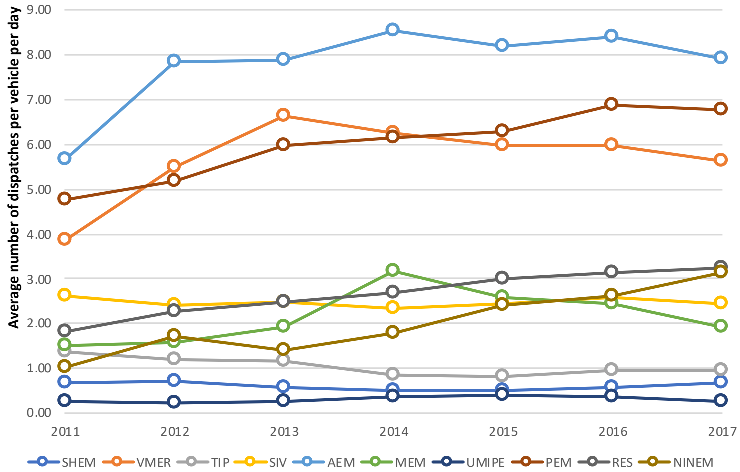


Figure 7 - Vehicle dispatches per day per vehicle. Source: *Ministério da Saúde*, 2018.

However, by considering Figure 8, which displays total number of dispatches per day (PEMs are plotted on the right axis), it is concluded that PEMs are the most requested vehicle, followed by RES and AEM, which are identical. This result could be expected given that, as mentioned, 75% of the fleet is PEMs.

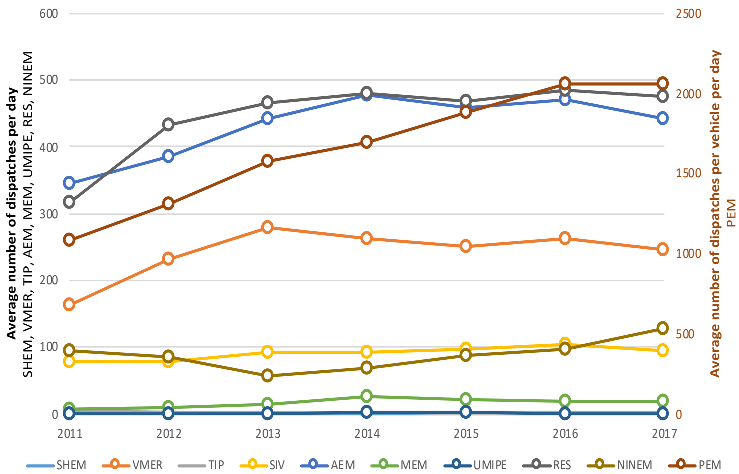


Figure 8 - Vehicle dispatches per day. Source: *Ministério da Saúde*, 2018.

After selecting the required vehicle type(s), the TEPH must dispatch actual vehicles on the field. For this purpose, a software displaying the locations of the emergency and nearby vehicles (whether stationed or in-transit, made available by a system called SIADEM) is used. The software presents a list of vehicles according to their proximity to the scene and a priority rule defined by INEM. Whenever possible, the TEPH dispatches INEM's vehicles first. If no INEM vehicle is available, a partner's vehicle is selected (i.e. a PEM or RES). Non-INEM vehicles (NINEM) are usually dispatched only when other ambulances (AEMs, PEMs or RESs) are unavailable/inexistent. Since there is no formal protocol establishing cooperation between INEM and the owner/operator of a NINEM, INEM pays a higher price when these vehicles are dispatched. Therefore, the software first presents AEM ambulances ordered by proximity, then PEM, RES and finally NINEM. A 15-kilometer isochrone around the emergency location is also presented. The TEPH dispatches the closest available vehicle – thus following the closest-available vehicle policy common in many EMS systems (Reuter-Oppermann, Van Den Berg and Vile, 2017; Bélanger, Ruiz and Soriano, 2018) – while also taking into account the priority order of vehicles. A vehicle is said to be idle if it is stationed at or traveling to a base or leaving a hospital after transferring a patient. In severe cases, a vehicle may be diverted from another less-urgent call.

Vehicles are typically dispatched to occurrences within their usual area of activity, which may be a municipality or a parish. This is especially true for PEM and RES ambulances, which are operated by local firefighters that are better acquainted with the territory. In exceptional situations, they may be dispatched to other zones if the severity of the emergency and the location of other vehicles justify it. RES ambulances provide a second level of resources, and they are intended to complement AEM and PEM ambulances (Instituto Nacional de Emergência Médica, 2017c).

The communication channel used for dispatching depends on the vehicle itself. For vehicles owned and operated by INEM, an integrated software called iCARE is used, conveying information about the call automatically to the crew. Nevertheless, telephonic confirmation is always required to ensure proper understanding of the occurrence. For other vehicles, the telephone or radio (including SIRESP, the national emergency communication network) are used. Having numerous communication channels introduces redundancy, allowing communication even when some of these systems are down.

3.2.5 ARTICULATION AND COORDINATION

A separate team of TEHPs at the CODU is responsible for coordinating all entities which respond to an emergency and for counselling emergency teams. They provide information about the victim's condition, ensure that adequate decision algorithms are followed and articulate the patient's arrival to the hospital. This team of TEHPs performs a major role at the CODU (Gomes *et al.*, 2004). Additionally, a team of regulating doctors is essential to provide fully integrated care (Instituto Nacional de Emergência Médica, 2015). Figure 9 summarizes the preceding sections about the SIEM.

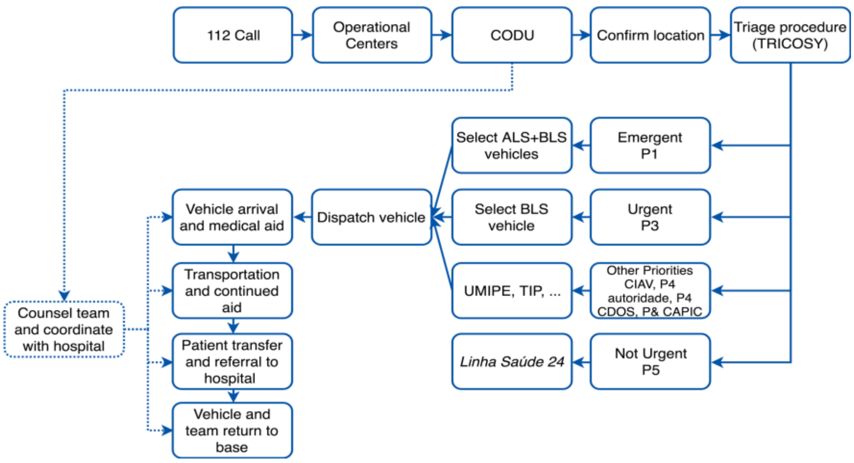


Figure 9 - SIEM's process

3.3 VEHICLE LOCATION PLANNING AT INEM

Response times are a critical factor in any EMS system. As such, selecting the location of emergency vehicles is an important decision. Although these decisions are taken in a dynamic environment, effective *a priori* planning can be beneficial (Goldberg, 2004).

Vehicle location planning is under the supervision of the GPCG and is the focus of this study. In this section, suitable vehicle bases, according to Portuguese Legislation and INEM's criteria, are introduced. Subsequently, the current short and long-term vehicle planning processes are detailed. Finally, the performance indicators used by INEM to evaluate vehicle positions are presented.

The GPCG is responsible for supporting several other planning decisions at INEM. For instance, it is responsible for scheduling TEPHs working at the CODU and operating INEM's vehicles. This task involves a time-consuming process, partially supported by a software called "*Gestão de Horários*", which fails to meet several requirements, such as legal constraints, TEPHs preferences and demand constraints (Rosa, 2017). It is also responsible for sizing the workforce of TEPHs, at the CODUs. Currently, this planning process is based on experience. The GPCG empirically forecasts demand based on past values and, subsequently, the historical ratio of calls answered by each TEPH at the CODU is used to estimate the necessary TEPHs. In locating emergency vehicles, the GPCP also follows an empirical procedure, described in the following sections.

3.3.1 VEHICLE STATION

A vehicle station is an infrastructure where an emergency vehicle begins and ends service and, in the case of INEM, is stationed while waiting calls. Although an ambulance is usually assigned to a single base, it is possible to reposition ambulances throughout the day in order to improve coverage, a practice called ambulance relocation. Therefore, although the position of an ambulance base is fixed, ambulances themselves are not. In the case of INEM, however, ambulance relocation is not performed. Whenever a vehicle finishes responding to an emergency, it returns to its original base station, where it replenishes its stock of medical supplies.

For most vehicles, the Portuguese legislation establishes rules for potential base locations, with the main goal of articulating the system with the network of emergency departments of NHS hospitals:

- VMERs and SIVs require differentiated staff and equipment, so they have to be located at emergency departments of NHS hospitals. These vehicles' crews are required, by law, to be integrated in the emergency teams of these hospitals. In Portugal, there are three types of emergency departments: Multipurpose Urgency Services (SUP, which handle basic emergencies), Medical-Surgical Urgency Services (SUMC, which attend emergencies requiring medical treatment or surgery) and Basic Urgency Services (SUB, which handle emergencies requiring differentiated treatment). All SUP and SUMC hospitals should theoretically be assigned a VMER, while SUB hospitals should have a SIV (*Despacho* 5561/2014);
- TIP ambulances must be located in all hospitals operating pediatric or neonatal intensive care units (*Despacho* 1393/2013);
- AEMs and MEMs must be located in regions where SUP or SUMCS hospitals exist (*Despacho* 10109/2014). Potential bases for these vehicles include hospitals, health centres, firefighters or police stations as long as they fulfil certain conditions. For this reason, the location of these vehicles is more flexible. Nevertheless, INEM must establish collaboration protocols to use these facilities as stations, since it is necessary to pay overhead costs. Additionally, since there are few MEMs, their location usually remains unchanged;

- PEMs are located, by definition, in firefighters or Red Cross facilities. The Ministry of Health has established that there should be at least one PEM in each municipality of the country (*Despacho* 10109/2014), in an attempt to promote equity and ensure territorial coverage;
- UMIPEs are located at the headquarters of INEM's Regional Delegations.

Therefore, INEM does not build dedicated ambulance stations. Instead, it uses existing facilities as bases, provided that they have enough space and conditions to be used.

3.3.2 VEHICLE LOCATION PLANNING PROCESS

Vehicle location planning at INEM comprises two processes: long-term and short-term planning. Long-term planning concerns choosing stations and assigning vehicles to those stations for everyday operation. This is an incremental process: whenever a new vehicle is to be added to the fleet, the GPCG helps the board decide where it should be positioned. Sometimes, CODU supervisors identify critical regions, where TEPHs consistently face difficulties in finding an appropriate vehicle. In that case, they suggest that region as a good candidate for a new vehicle, which is further analysed by the GPCG.

Therefore, location planning is not approached holistically, as INEM does not change the location of already positioned vehicles in the long run. For example, although it would be legally possible to reassign a PEM ambulance to a different firefighter base, this option is not well-viewed by the GPCG.

In urban areas, INEM divides the territory into parishes, while municipalities are used in rural areas. In order to choose the area for a new vehicle, the GPCG evaluates demographics (population density) and emergency history. It also considers the region's existing vehicles. After empirically gauging this demand/supply balance, the GPCG uses experience to choose where the vehicle is more beneficial.

Subsequently, a station within that area must be selected. Obviously, the constraints mentioned above should be respected. In practice, this is not an easy task, as the unavailability of suitable locations (according to the Portuguese legislation) may require a certain area to be less covered than desirable. For example, as of 2016, 8 out of 39 SIV ambulances were not integrated in the SUB emergency departments of hospitals because it was impossible to find a suitable location that covered the territory appropriately. VMERs, however, were fully integrated (Instituto Nacional de Emergência Médica, 2018).

Figure 10 shows the location of existing vehicles. AEM ambulances and MEMs are located mainly on the coast, near urban areas, whereas PEM ambulances cover the rest of the territory, but also tend to be more concentrated on urban areas.

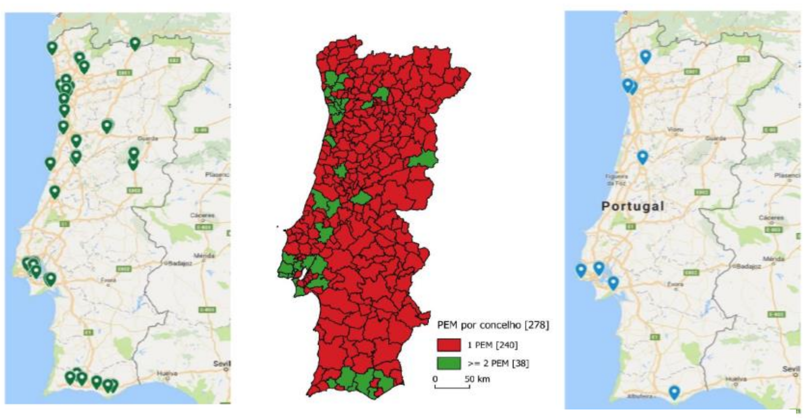


Figure 10 - Locations of AEM (left), PEM (middle) and MEM (right) vehicles in Portugal. Source: Instituto Nacional de Emergência Médica, 2017c

Short-term planning concerns temporarily repositioning vehicles in face of disruptive events. In this situation, INEM looks at existing vehicles and decides which adjustments are required to increase the number of vehicles in the target area while ensuring that the remaining territory is not uncovered. The disruptive event can be expected or unexpected. If it is expected, vehicles must be positioned in stand-by locations closer to critical areas. Here, forecasting based on experience is used to determine where the likelihood of an emergency is higher. INEM can use back-up vehicles or reposition active ones. For instance, in Lisbon, back-up vehicles are kept in *Lumiar*. In practice, activating back-up vehicles is rare due to crew unavailability, high activation costs and bureaucratic approvals. Hence, it is common to relocate active vehicles, specifically AEMs, since PEMs require special authorizations. For example, during the 2014 Champions League Final in Lisbon, an AEM from *Sacavém* was repositioned closer to the stadium. If there is an unexpected event, the GPCG relocates nearby vehicles to the area first and then decides on how to reposition the remaining vehicles to cover areas exposed by the first relocation.

3.3.3 PERFORMANCE MEASURES FOR VEHICLE LOCATION PLANNING

To evaluate the system's performance regarding vehicle location, identify improvement opportunities and support location decisions, the GPCG uses three main performance indicators:

- **Rescue Time**: time between vehicle dispatch and arrival, measures the proximity of vehicles to the emergency. This is different from Response Time, which is the time between the call being received and arrival on scene. INEM targets rescue times of less than 15 minutes in urban areas and 30 minutes in rural areas. Rescue time data are available for AEM and SIV ambulances (Figure 11), showing that these targets are not always met. For the remaining INEM vehicles, this data is can be retrieved from emergency logs. For RES and NINEMs, there are no data;

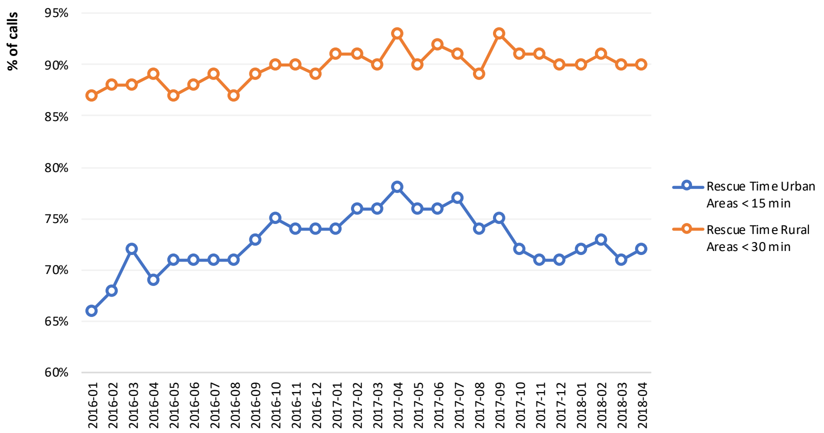


Figure 11 - Rescue times of AEM and SIV ambulances since 2016. Source: Instituto Nacional de Emergência Médica, 2018a.

- **External Dependency**: fraction of emergency calls which are answered by vehicles from neighbouring regions. If this proportion is significative, the region is not self-sufficient, meaning that additional vehicles are required, or existing ones must be repositioned;
- **Dispatch Time**: time between defining the priority of an emergency and the TEPH dispatching the appropriate vehicle. This is an indicator of the closeness of vehicles to the emergency. If a vehicle is nearby, the TEPH automatically dispatches it. Otherwise, the TEPH needs to search further away, taking more time. The main advantage of this indicator is that it is easily measurable, but it is only a proxy for the proximity of available vehicles to the emergency scene.

As stated, the Response Time is the most common performance measure of EMS systems, and some countries' legislation establishes specific targets (Reuter-Oppermann, Van Den Berg and Vile, 2017). For instance, UK's EMS must reach 90% of Category 1 (similar to P1) calls in under 8 minutes. Portuguese legislation, however, does not set any target regarding response time nor rescue time.

3.4 PROBLEM DEFINITION

The aim of this dissertation is to develop and apply a mathematical model to support vehicle location planning at INEM. The model should aid INEM in positioning the emergency fleet, identify weak spots in the territory, test the impact of different legal and self-imposed constraints and set goals on the performance of the existing system. The model must take into account some restrictions, such as limited resources, multiple types of vehicles, legal restrictions on suitable bases and multiple care providers with different levels of expertise. The model will be applied to the cities of Lisbon and Setúbal, focusing on P1 and P3 calls, which account for the majority of emergencies, and corresponding response vehicles: SIV, VMER, AEM, PEM, RES and NINEM.

Lisbon and Setúbal are two different, yet some of the most challenging, urban areas in what concerns emergency vehicle location planning. For this reason, they have been selected, in articulation with the GPCG, as case-study areas for this dissertation.

Lisbon is the capital and the mostly populated municipality in the country. Emergency care is provided mainly through AEM vehicles, complemented by few PEM and RES ambulances, three VMERs and one SIV. Therefore, it is a region where vehicle location planning is more flexible and, simultaneously, more pressured. Setúbal, on the other hand, relies more on PEM ambulances to cover the territory. Furthermore, it is frequently necessary to dispatch vehicles from neighbouring regions due to the unavailability of its own vehicles. Since its External Dependency is high, the GPCG believes that there are not enough vehicles in the region or, alternatively, that their positioning could be improved. Besides, this region has never been explored in previous studies of the same nature.

3.5 CHAPTER CONCLUSIONS

INEM plays a fundamental role in delivering emergency care in Portugal. It coordinates and operates a complex EMS system comprising multiple stages, entities, types of emergencies and vehicles. Naturally, planning such a system becomes challenging and although it is recognized that proper vehicle planning is paramount for EMS effectiveness (Goldberg, 2004), INEM relies mostly on experience and intuition to make key decisions, including the location of emergency vehicles.

The goal of dissertation is to create a mathematical model to assist INEM and, specifically, the GPCG, in locating emergency vehicles and, subsequently, achieving a better service level with the available resources. This model should give solid insight into the impact of different vehicle configurations and contribute to the development of vehicle location plans. At this stage, it is important to review the existing literature on EMS vehicle location planning, in order to understand what approaches can be used to address this problem. This information is useful to support the development of a methodology to tackle the problem and is the focus of Chapter 4.

4. LITERATURE REVIEW

This chapter reviews existing literature concerning the EMS vehicle location problem. After a general overview of the facility location problem and the EMS location literature, selected models are reviewed. Solution techniques are presented alongside each model. Sections 4.1 and 4.2 introduce the facility location problem and present an overview of EMS location models. Sections 4.3 to 4.9 present selected optimization models. Section 4.10 addresses other issues, including the role of Simulation. Finally, section 4.11 presents the chapter's conclusions. A summary table is provided in Appendix B.

4.1 THE FACILITY LOCATION PROBLEM

Location decisions arise in public and private sectors and are usually long-term investment decisions (Daskin, 1995; Owen and Daskin, 1998). Therefore, building location models has been a topic of OR research for many years, under the field of Location Science (ReVelle, Eiselt and Daskin, 2008).

A facility can be a warehouse, a school, a factory or an ambulance station. It is built to meet demand, which for EMS is emergency requests (ReVelle *et al.*, 1977; ReVelle, Eiselt and Daskin, 2008). Location models seek to answer questions as how many facilities to operate, where to locate them and how to use them to meet demand (Daskin, 1995). As any Mathematical Programming model, they include decision variables, input parameters, objective function(s) and constraints.

ReVelle et al. (2005) highlight four key components of all location models:

- **Customers** generating demand, whose locations are known. These locations are often aggregated in discrete points, for computational simplicity (Li, et al., 2011);
- **Facilities** to be located, usually consisting of a set of existing facilities and potential new facility sites (Melo, Nickel and Saldanha-da-Gama, 2009);
- A **space** where facilities and customers are located. Most models consider a discrete space but continuous models also exist (Farahani *et al.*, 2012). This review concerns discrete models;
- A **metric** measuring time, distance or cost between facilities and clients.

Solving these models yields an optimal location plan, which maximizes (or minimizes) an objective function. Naturally, the location plan proposed by the model is only "optimal" as long as its assumptions are valid (Goldberg, 2004). Additionally, if there are multiple conflicting objectives, no optimal location plan may exist.

Location models therefore require a quantifiable description of the decision-maker's objectives (Daskin, 1995). It is generally recognized that private-sector objectives are easier to state than public-sector ones. In particular, the goal of EMS is usually expressed as the maximization of public benefits or, conversely, the minimization of losses (Indriasari *et al.*, 2010).

Naturally, there are many different types of location models. Nevertheless, these can be classified into 5 basic categories, according to the problem they address and their objective (Laporte and Nickel, 2015):

- **P-median models**: locating *p* facilities to minimize the weighted distance between customers and the nearest facility;
- **P-centre models**: locating *p* facilities to minimize the maximum distance between any customer and its closest facility;

- **Covering models**: locating facilities to maximize covered customers, i.e. those within a certain distance/time standard of the closest facility;
- **Anti-covering models**: locating repulsive facilities, seeking to minimize population coverage;
- **Fixed-charge models**: locating facilities and allocating them to demand to minimize fixed-charged opening costs and assignment costs while respecting capacity constraints.

4.2 THE EMS LOCATION LITERATURE

EMS facility location is one of the most exciting areas within facility location since the 1960s (Goldberg, 2004). In fact, many models were developed to address EMS location problems and only latter used in different settings, perhaps due to the high social impact of these systems (Aringhieri *et al.*, 2017).

Consequently, the EMS location literature is rich, and many types of location models have been proposed to address this issue (Başar, Çatay and Ünlüyurt, 2012). Nevertheless, covering models are prevalent because EMS are usually evaluated on coverage metrics (Reuter-Oppermann, Van Den Berg and Vile, 2017).

Location models are usually Integer (IP), Binary Integer (BIP), Mixed Integer (MIP) or even Non-Linear (NLP) Programming models. Since solving real-sized instances can be challenging, different solution approaches have been proposed, including heuristics, metaheuristics and exact techniques, and this is a topic drawing substantial research (Owen, et al., 1998). These heuristics seek to find good solutions in short computational times by exploring the underlying formulations (Daskin, 1995).

This chapter focuses on discrete location models for strategic and tactical planning. Much recent research has focused on operational decisions of relocation and dispatching (Bélanger, Ruiz and Soriano, 2018). Since operational issues are not the focus of this study, these models are excluded.

Besides optimization models, two other classes of OR models are popular in the EMS literature: Simulation and Queueing Theory. This review focuses only on analytical optimization models, because their prescriptive nature allows us to produce recommendations regarding the location of emergency vehicles. Therefore, the use of Queueing Theory and Simulation with optimization models is mentioned, but not reviewed in detail herein.

A particularly important queueing model is the Hypercube Queueing Model (Larson, 1975), which models a spatially-distributed system in which the state of each vehicle is tracked. This model is suited for server-to-customer queueing systems and it allows many important performance metrics to be derived from the state probabilities (that is, the probability that different combinations of vehicles are busy) including average response times, fractions of dispatches and server workload.

The following sections proceed as follows. After presenting two seminal covering models, key aspects of EMS operation which have been added to the original formulations, including vehicle unavailability, demand and travel time uncertainty, multiple vehicle types and call priorities, decision-making objectives, time-dependency and integrating multiple decision levels are introduced. A brief review of simulation-optimization and other issues, such as forecasting, demand aggregation and model validation, is also presented.

4.3 EARLY WORKS: STATIC DETERMINISTIC COVERING MODELS

Covering models explore the notion of demand coverage: a demand node is covered if it can be reached by at least one facility within a certain distance/time limit, called coverage radius (Toregas *et al.*, 1971).

The first covering model was the Location Set Covering Problem (LSCP), introduced by Toregas et al. (1971). The LSCP determines the minimum number and location of facilities that guarantee coverage of all demand nodes. The authors propose solving the Linear Relaxation (LR) and employing a single cut to remove fractional results as solution techniques. Branch-and-Bound (BB) and Reduction Techniques were also proposed (Church and Meadows, 1979).

The LSCP is attractive due to its simplicity and accuracy in capturing real EMS needs (ReVelle *et al.*, 1977). Unfortunately, a major drawback is that covering all demand nodes may require a number of facilities that are not available. In reality, decision-makers may be more interested in knowing what can be achieved with the available limited resources (Bélanger, Ruiz and Soriano, 2018).

Therefore, Church and ReVelle (1974) introduced the Maximum Covering Location Problem (MCLP), which maximizes demand coverage by locating a limited number of facilities. The MCLP can be used to select among optimal solutions of the LSCP and to derive cost-effectiveness curves by varying the number of available vehicles. To solve real-sized instances, two greedy heuristic procedures, solving the LR with inspection and BB are proposed. This model was used by Eaton *et al.* (1985) in a study in Austin, Texas, demonstrating how the MCLP can be used to study multiple objectives, stakeholder opinions and vehicles only by changing input parameters.

Both these models were very influential and have inspired most recent approaches (Farahani *et al.*, 2012). However, they relied on strong assumptions, such as unlimited vehicle availability, a single vehicle type, deterministic demand, travel time and service time (Goldberg, 2004). They also employed a simple coverage definition – a demand node is either fully covered or not – although, in reality, service quality degrades as distance increases (Murray, 2016). Since incorporating uncertainty provides more realistic solutions (Erkut *et al.*, 2009), researchers have been extending the original formulations to get accurate representations of EMS systems. In the sections bellow, approaches to add key EMS components to the original works, starting with vehicle unavailability, are presented.

4.4 DEALING WITH VEHICLE UNAVAILABILITY

Initially, researchers focused on the assumption of unlimited vehicle availability. Obviously, when a vehicle is dispatched, its designated regions are no longer covered. Thus, the solutions proposed by the LSCP and MCLP may not perform well enough for real-world applications (Bélanger, Ruiz and Soriano, 2018). To deal with this issue, three alternatives have been explored: multiple coverage models, probabilistic models and capacitated models (Li *et al.*, 2011).

4.4.1 MULTIPLE COVERAGE MODELS

Multiple coverage models cope with the unavailability of vehicles by ensuring that more than one vehicle is capable of covering a demand node. Daskin and Stern (1981) were the first using this approach, extending the LSCP by adding the secondary objective of maximizing nodes with back-up coverage. Later, Hogan and ReVelle (1986) proposed two Backup Coverage Models (BACOP 1 and 2), maximizing the population covered twice while ensuring basic coverage and studying the trade-off between these two goals, respectively. The same trade-off was also studied by Storbeck (1982) using Goal Programming.

Recognizing that double coverage may not be appropriate for all nodes, Batta and Mannur (1990), Church and Gerrard (2003) and Degel *et al.* (2015) provided different coverage levels to each node. These can be set by analysing the probability distribution of concurrent demands (Degel *et al.*, 2015)

Yet, the most influential multiple coverage model was the Double Standard Model (DSM). The DSM maximizes back-up coverage, while ensuring basic coverage of a fraction of demand and full coverage within a larger coverage radius (Gendreau, Laporte and Semet, 1997).

The DSM has been frequently applied to real problems (Laporte *et al.*, 2010): in Vienna, Austria, (Doerner *et al.*, 2005); in Shanghai, China (Su, Luo and Huang, 2015) and in Chicago, USA (Liu *et al.*, 2016). Solution techniques for the DSM include Tabu Search (Gendreau, Laporte and Semet, 1997), Ant Colony (Doerner *et al.*, 2005) and Genetic Algorithm (GA) (Liu *et al.*, 2016).

Multiple coverage models have the advantage of being easy to understand, but do not explicitly ensure that a vehicle is available. Besides, double coverage may not actually be required for all nodes.

4.4.2 PROBABILISTIC MODELS

A different approach is to consider explicitly the unavailability of emergency vehicles through a probabilistic parameter called busy fraction. A vehicle's busy fraction is the probability or fraction of time that it is unavailable to serve demand (Bélanger, Ruiz and Soriano, 2018). By using this parameter, it is possible to account only for the real availability of a vehicle. These models differ on three main areas: the scope of the busy fraction, the approach used to calculate it and how it is used in the model.

Regarding the busy fraction's scope, an alternative is to use a system-wide busy fraction (i.e. equal for all vehicles), independent of their locations (Daskin, 1983). By dividing the region in independent demand subareas, area-specific busy fractions can be calculated, being equal for all vehicles in that subarea (ReVelle and Hogan, 1988). Closer to reality, server-specific busy fractions require treating this parameter as endogenous, leading to computationally difficult formulations (Goldberg and Paz, 1991).

To calculate the busy fraction, the initial approach was to use historical data to calculate a system-wide busy fraction by dividing total time spent serving calls and total vehicle availability time. To obtain more accurate estimates, Simulation (e.g. (Davis, 1981)) and the Hypercube Queueing Model (HQM) proposed by Larson (1974) have been frequently used within optimization models. The advantage of HQM-based models is that they are analytical and, hence, usually quicker to solve (Ingolfsson, 2013).

Finally, models differ on the way they explore the busy fraction. Three approaches can be identified: reliability models, expected coverage models and hybrid models. Reliability models establish that a node is covered if the probability that one vehicle is available within the coverage radius is greater than a pre-specified level (ReVelle and Hogan, 1988). Expected coverage models maximize the expected population receiving appropriate service by emergency vehicles, given by the probability that each vehicle is available times the demand it covers (Daskin, 1983). Hybrid models mix both approaches.

Selected probabilistic models are presented next, starting with reliability and expected coverage models. Subsequently, hybrid approaches are discussed. Queuing-based versions of both classes of models are considered next, followed by models with server-specific busy fractions.

Reliability Models

Chapman and White (1974) were among the first to formulate a reliability model with their Probabilistic LSCP (PLSCP). The PLSCP assumes a known uniform busy fraction, although Simulation is proposed to obtain better estimates. Later, ReVelle and Hogan (1988) extended this model to include area-specific busy fractions and reformulated the reliability constraint to derive the minimum number of vehicles required to cover each area with the desired reliability. In subsequent papers, the authors introduce variations: the α-Reliable p-Center Problem, the Maximum Reliability Location Problem and, more important, the Maximum Availability Location Problem (MALP). The MALP locates facilities to maximize covered population with α-reliability, using either system-wide (MALP1) or area-specific (MALP2) busy fractions (ReVelle and Hogan, 1989b, 1989a). Application of MALP models shows that using area-specific instead of system-wide busy fractions leads to more distributed solutions. It is recognized, however, that server-specific busy fractions would be more desirable.

The REL-P model, formulated by Ball and Lin (1993), finds the minimum cost vehicle configuration that ensures coverage of all nodes with a given reliability, using area-specific busy fractions and dependent servers. ReVelle and Marianov (1991) also applied chance-constraints to create the probabilistic FLEET (PROFLEET), maximizing the population covered by two vehicle types with the same reliability level.

Most of these models, however, assume server independence – i.e. the probability of one vehicle being busy does not affect the probability of other vehicles being busy (Marianov and ReVelle, 1996). Additionally, Erkut (2008) argued that the objective function of reliability models does not reflect the performance measures used by EMS practitioners. Reliability models can have unintended consequences, because a node that is covered with a given reliability contributes fully to the objective function, while a node covered with a slightly lower reliability does not (Sorensen and Church, 2010). In this sense, expected coverage models may be more appropriate (Erkut, 2008).

Expected Coverage Models

The first expected coverage model was the Maximum Expected Coverage Location Model (MEXCLP), introduced by Daskin (1983). The model assumed a known system-wide busy fraction. Each additional ambulance within the coverage radius marginally increases the expected coverage of a node, as it is useful when others are busy. Thus, locating more than one ambulance at the same base can be beneficial. To obtain the increment in expected coverage provided by an ambulance, one must multiply demand by the probabilities that the remaining vehicles are busy, and that particular vehicle is not. Since the MEXCL is a large IP model, a single-node substitution heuristic was originally proposed.

Several variations of the MEXCLP have been subsequently proposed: combinations with the DSM (Chuang and Lin, 2007); multiple facilities and equipment with different busy fractions (Bianchi and Church, 1988; Jayaraman and Srivastava, 1995); time-dependent versions (Repede and Bernardo, 1994) and a multi-vehicle version - the MOFLEET model (Bianchi and Church, 1988).

The MEXCLP and its variants provide a suited objective for EMS location. However, they mostly rely on the unrealistic assumption that vehicles operate independently, with the same known busy fraction regardless of their locations (Batta, June M Dolan and Krishnamurthy, 1989). This assumption does not necessarily hold in practice, as vehicles in different regions may have significantly different workloads.

Hybrid Approaches

Some authors have successively combined reliability with expected coverage. Alsalloum and Rand (2006) proposed a Goal Programming model: the first goal was to locate stations to maximize coverage probability, while the second is to allocate ambulances to each station to minimize spare workload while ensuring a busy fraction bellow 5% (in essence, a reliability constraint). An analysis to each demand area is conducted using Erlang's Loss Formula to ensure that the adequate number of ambulances are located. By disaggregating the model into two stages (one for each goal), optimal solutions are efficiently obtained. Sorensen and Church (2010) combined the local reliability and area-specific busy fractions of the MALP with the MEXCLP's objective. Optimal solutions based on data from Washington, D.C., are obtained using CPLEX. Simulation shows that the original reliability models have unsuited objectives but confirm that local busy fractions provide more accurate descriptions of the system.

Queueing-based Models

To overcome the assumptions of previous models in calculating busy fractions, queueing models have been embedded in optimization models. Here, the HQM (Larson, 1974, 1975) and the Approximate Hypercube Queueing Model (AHQM) (Jarvis, 1985) are typical choices.

Marianov and ReVelle (1994, 1996) proposed the Queueing PLSCP (Q-PLSCP) and Queuing MALP. Both models conceptualize each subarea as an isolated $M/G/s-loss$ queuing system. Consequently, they calculate area-specific busy fractions and arrive at the number of vehicles required to cover each area. In turn, Marianov and Serra (1998) and Amiri (1998, 2001) conceptualized each facility as a queueing system and developed models including limits or penalties on customer waiting time.

Batta, Dolan and Krishnamurthy (1989) formulated the Adjusted MEXCLP (AMEXCLP) and a Hypercube Optimization Procedure. The AMEXCLP introduced corrective factors coming from the HQM in the MEXCLP's objective function which relax the server independence assumption. Correction factors have since been extensively used (e.g. (Ingolfsson, Budge and Erkut, 2008; Rajagopalan, Saydam and Xiao, 2008; McLay, 2009)). The Hypercube Optimization Procedure used expected coverage calculated from the AHQM to guide a single-node substitution heuristic. An analogous approach was proposed by Saydam and Aytuğ (2003), who embed an AHQM into a GA, and Galvão, Chiyoshi and Morabito (2005), who use correction factors from the AHQM to calculate busy fractions and apply a Simulated Annealing (SA) algorithm, based on a single node substitution heuristic.

Site-Specific Busy Fractions

Most models so far treat the busy fraction as an exogenous parameter, although, in reality, it is consequence of the chosen location pattern (Aringhieri *et al.*, 2017). This simplification is mainly due to the size and nonlinearity of models using endogenous busy fractions, resulting from their multiplication in the objective function or in chance-constraints (Cho *et al.*, 2014).

One approach to calculate site-specific busy fractions is to use iterative procedures that calculate the busy-fraction for each potential solution. Knight, Harper and Smith (2012) and ReVelle and Hogan (1988) both tested this approach, but failed to reach a stable solution. Inversely, Ingolfsson, Budge and Erkut (2008) successfully iterated between the optimization model and the HQM.

Other authors calculate site-specific busy fractions directly inside the optimization model. To do so, standard Queueing Theory or the HQM can be used. For instance, Borras and Pastor (2002) adjusted the REL-P and PLSCP to include server-specific busy fractions using the queueing system of Marianov and ReVelle (1994, 1996). Toro-Díaz *et al.* (2013) included the balance equations of the HQM directly into the constraints of their model. However, this formulation is too challenging to be solved exactly.

Alternatively, the busy fraction can be estimated by dividing the expected worktime for each vehicle by its availability time, according to the chosen location. Goldberg and Paz (1991) calculated server-specific busy fractions by establishing ordered rankings of dispatching preference and formulating non-linear equations. The resulting model is a NLP model, leading to a single and double-node interchange heuristics to generate solutions in reasonable time. Cho *et al.* (2014) explicitly assigned patients to rescue helicopters to derive the busy fraction, while Shariat-Mohaymany *et al.* (2012) developed a linear model which calculated ambulance workload by assuming that emergency requests are distributed evenly among ambulances within the coverage radius.

Finally, Leknes *et al.* (2017) implicitly determined busy fractions by allocating demand to primary and secondary stations and calculating the resulting arrival and service rates. Queuing Theory results are used and linearized using Special Ordered Sets. The model was applied using commercial solvers.

4.4.3 CAPACITATED MODELS

Another approach to deal with vehicle unavailability consists in limiting the demand that each vehicle can cover, i.e. establishing capacity constraints. These models are called Capacitated Models.

Initially, Current and Storbeck (1988) formulated a capacitated MCLP considering only demand within the coverage radius for capacity constraints. This model was expanded by Pirkul and Schilling (1991) to include demand inside and outside the coverage radius, resulting in a challenging formulation and subsequent development of an heuristic. Capacity constraints are also used by Shiah and Chen (2007) and Schmid and Doerner (2010). All these models use a uniform capacity limit on all facilities. To overcome this limitation, Yin and Mu (2012) formulated a model with varying capacity levels for each base according to its number of vehicles.

Capacitated models are simpler to understand than probabilistic versions and may be suited for high-level planning (strategic). Nevertheless, determining the appropriate capacity limits may be difficult and the congestion effects on the system may be overlooked.

4.5 MODELLING STOCHASTIC DEMAND AND TRAVEL TIME

Demand and travel time are important sources of uncertainty in EMS systems and have naturally gained the attention of many researches. Demand is uncertain since it can arise at any time and at any location. Travel time depends on the caller's location and on traffic conditions (Bélanger, Ruiz and Soriano, 2018). In this section, approaches to model these factors are covered, starting with demand followed by travel time, while Fuzzy Programming, which can be used to model both, is discussed at the end.

4.5.1 DEMAND

There are several methods for modelling uncertain EMS demand. Models using Queueing Theory usually assume that demand follows a Poisson process so that standard Queueing Theory results can be used (Batta, June M. Dolan and Krishnamurthy, 1989; Marianov and ReVelle, 1994, 1996; Marianov and Serra, 1998; Borras and Pastor, 2002; Saydam and Aytuğ, 2003; Galvão, Chiyoshi and Morabito, 2005; Cho *et al.*, 2014; Leknes *et al.*, 2017). Similarly, the REL-P models demand as a Poisson process.

Demand can also be treated as a random variable with general probability distributions, which can be used to write chance-constraints. For instance, Beraldi, Bruni and Conforti (2004) developed a stochastic model (and its deterministic counterpart) which minimized costs subject to global reliability demand satisfaction constraints. Zhang and Li (2015) presented a chance-constrained model to ensure that the existing vehicles can handle the maximum simultaneous demands with a given probability. These approaches require demand to be described by a known probability distribution. In practice, this may not be possible due to data unavailability. For this reason, Chu *et al.* (2018) proposed a distribution-free chance-constrained model, in which reliability is guaranteed for all possible probability distributions with the estimated parameters (mean and covariance).

An alternative approach is to use demand scenarios. Beraldi and Bruni (2009) presented a two-stage Stochastic Programming model in which the first-stage decision selected bases and their capacities. After uncertainty is resolved, vehicles are assigned to bases and demand requests are assigned to vehicles. Zhang and Jiang (2014) developed a bi-objective robust location-allocation model which minimizes costs and unsatisfied demand while ensuring minimum levels of demand satisfaction for a range of demand realizations (number of calls and number of concurrent calls). The same approach is used by Nickel, Reuter-Oppermann and Saldanha-da-Gama (2016) who, realizing the intractability for larger instances, proposed a demand sampling approach. The idea is to sample a subset of scenarios which allow the model to be solved. Using the optimal solutions of each sampling iteration, the optimal solution of the whole problem can be approximated. Sung and Lee (2018) formulated a Stochastic Programming model using sampled demand scenarios considering the temporal order of calls. In the first stage, the model locates ambulances to maximize expected coverage over all scenarios. In the recourse decision, ambulances are assigned to calls depending the dispatching rule, provide service and return base, becoming available. This approach allows the availability of ambulances to be explicitly considered. Given the difficulty in solving Integer Stochastic Programming models, a Benders decomposition algorithm is proposed. Berman, Hajizadeh and Krass (2013) developed three models which consider different demand and travel time scenarios. A Greedy heuristic and a Lagrangian Relaxation heuristic are proposed to solve the model efficiently. Finally, Bertsimas and Ng (2019) use structured uncertainty sets to model demand interactions across heterogeneous regions and use constraint-generation to solve the corresponding stochastic and robust two-stage models.

4.5.2 TRAVEL TIMES

Similar approaches can be used to model travel time uncertainty. Scenarios were used, for instance, by Berman, Hajizadeh and Krass (2013). Differently, Aly and White (1978) treated the location of an emergency as a random variable uniformly distributed over a region. In the equivalent deterministic model, a facility can cover a node only if the probability that travel time is lower than the coverage radius is above a threshold. Marianov and ReVelle (1996) modelled travel times as normal variables and recalculate the coverage areas of each facility using this information.

A different approach is to use the probability that a vehicle can reach a demand area within the coverage radius to determine the expected coverage (Daskin, 1987; Goldberg and Paz, 1991). The tail of an

Exponential Distribution is used to approximate the probability distribution, although the authors indicate that other distributions may be more suited. This implies that the probability of a vehicle covering a demand node depends on its availability and on the probability that the travel time is bellow a threshold. Ingolfsson, Budge and Erkut (2008) expanded this approach by accounting for the pre-travel delay, while Drezner, Marianov and Wesolowsky (2016) added the acceleration-deceleration effect.

4.5.3 FUZZY MODELS

Fuzzy programming is useful when quantitative data regarding uncertain parameters cannot be obtained, since it requires only qualitative judgements from the decision-maker. Therefore, they have been used to model demand uncertainty (Wen and Iwamura, 2008), travel time uncertainty (Davari, Fazel Zarandi and Hemmati, 2011; Lahijanian, Zarandi and Farahani, 2017) or both (Torres, Trujillo and Maldonado, 2018). These models introduce fuzzy triangular variables on traditional formulations, such as the MCLP (Davari, Fazel Zarandi and Hemmati, 2011) or the DSM (Lahijanian, Zarandi and Farahani, 2017). Usually, hybrid algorithms combining exact (e.g. BB, Goal Programming) and metaheuristic methods (e.g. SA, GA, Fuzzy Simulation) are used to solve these models efficiently.

4.6 TIME-DEPENDENT MODELS

The performance of EMS can change throughout the day (Matteson *et al.*, 2011), since traffic and demand conditions may leave some emergencies uncovered. Therefore, considering time-dependency can have a significant impact in the quality of a proposed location solution (Degel *et al.*, 2015)

Thus, many researchers transformed classical formulations into time-dependent models, including MEXCLP (Repede and Bernardo, 1994; Van Den Berg and Aardal, 2015), PLSCP (Rajagopalan, Saydam and Xiao, 2008; Setzler, Saydam and Park, 2009), MALP2 (Cheu, Lei and Aldouri, 2010), DSM (Schmid and Doerner, 2010; Dibene *et al.*, 2017) and BACOP1 (Başar, Çatay and Ünlüyurt, 2011).

In these models, travel time, demand, fleet size and station capacity can be used as dynamic parameters (Cheu, Lei and Aldouri, 2010). However, some authors do not take into account the cost of relocating vehicles and stations, thus leading to very volatile EMS systems with associated costs. This problem has been addressed in recent extensions (Van Den Berg and Aardal, 2015).

Additionally, Degel *et al.* (2015) provide a different coverage level for each time period and demand node, calculated by analysing the probability distribution of concurrent demands. Their model also includes travel times variations, relocation penalties and flexible fleet sizing. Tests show that considering time-dependency has a significant impact on the quality of the solution.

Since time-dependent formulations are more complex, alternative solution procedures have been proposed: Tabu Search (Saydam *et al.*, 2013); Variable Neighbourhood Search (VNS) (Schmid and Doerner, 2010) or commercial solvers (Dibene *et al.*, 2017).

These models usually split the day into time-periods. Thus, they are suited for operational decisions. However, studies considering multiple periods in the long-term planning horizon (i.e. months, years) are rarer. These approaches have been used in Supply Chain Network Design an existing network is transformed throughout the planning horizon (Melo, Nickel and Saldanha Da Gama, 2006, 2009). A possible approach is to modify location decision variables to account for different time periods and allowing capacity relocation at a cost, considering existing, selectable (i.e. those that can be relocated) and non-selectable facilities (i.e. those that cannot) (Melo, Nickel and Saldanha Da Gama, 2006).

4.7 MULTIPLE VEHICLES AND CALL PRIORITIES

Although early models include only one vehicle and emergency type, real systems have distinct vehicles and call priorities (Goldberg, 2004). In this section, approaches to model this feature are reviewed.

When dealing with multiple vehicles, a call can be covered if it is covered by any vehicle (Jayaraman and Srivastava, 1995; Coskun and Erol, 2010), by different vehicles simultaneously (David A. Schilling *et al.*, 1979) or even by multiple vehicles of each type. In some cases, different emergency types may require different vehicle combinations (Revelle and Snyder, 1995), with some vehicles being able to answer multiple call types (e.g. (Amiri, 1998)). The latter is the case of the SIEM, for instance.

Early multiple vehicle models include the TEAM (Tandem Equipment Allocation Model), FLEET (Facility-Location Equipment-Emplacement Technique) (Schilling *et al.*, 1979) and the FAST (Fire and Ambulance Service Technique) (Revelle and Snyder, 1995). These models maximized basic coverage of multiple types of calls, which required different vehicle types to be available within coverage radius.

Other researchers focus on the hierarchical nature of health systems. For example, Charnes and Storbeck (1980) proposed a Goal Programming model with three goals: critical call coverage with ALS, critical call back-up coverage with BLS and BLS coverage of non-critical calls. Serra (1996) formulated a model with two types of hierarchical facilities, ensuring that nodes covered by the same lower-level facility are also covered by the same higher-level facility.

Some models use variable coverage radius for different call priorities or place higher weights on covering critical calls. For instance, Liu *et al.* (2016) located ALS and BLS vehicles with different coverage radii according to the call's severity, while Van den Berg, Legemaate and Van Der Mei (2017) applied specific coverage radii for different vehicles, demand node and crew type combinations. Weights were also used by Colombo, Cordone and Lulli (2016), who presented a MCLP for multiple facilities and calls, alongside two Greedy Algorithms and a Heuristic Concentration Algorithm.

Other approaches to model differentiated call priorities use Queueing Theory. Silva and Serra (2008) proposed the Priority Q-LSCP in which each call priority has a different limit on waiting time. The model's complexity leads to the implementation of a Greedy Randomized Adaptive Search Procedure. McLay (2009) considered three call types in their two-vehicle MEXCLP. The goal is to maximize expected coverage of Priority 1 calls and the HQM is used to calculate busy fractions taking into account server interactions and lower priority calls. Similarly, Davoudpour, Mortaz and Hosseinijou (2014) considered different combinations of ALS and BLS vehicles that can be requested and use the HQM to estimate busy fractions, assuming a fixed-preference priority rule. Chong, Henderson and Lewis (2016) locate ALS and BLS vehicles to answer two types of calls, modelling interactions via a Markov Decision Process. They show that tiered EMS systems can be as good as ALS-only systems, and cheaper.

Although the problem of locating a heterogeneous emergency fleet has been addressed, a general framework for modelling multiple vehicles has not yet been achieved. Most of the existing models are designed specifically for the application at hand, having "hard-coded" constraints for the exact number and type of vehicles. As the SIEM operates multiple vehicles, a general framework would be desirable.

4.8 ALTERNATIVE PERFORMANCE MEASURES

Capturing the decision-maker's objectives is especially difficult in public decision-making. Many models so far consider a single coverage objective. However, combining various performance measures into location models may be beneficial (Knight, Harper and Smith, 2012). In this section, three alternative classes of performance measures are presented, as well as multi-objective approaches.

4.8.1 EXPANDING THE CONCEPT OF DEMAND COVERAGE

The original notion of coverage has three shortcomings: either a facility totally covers a demand node or not (all-or-nothing coverage); the coverage radius is an input, not a decision-variable; coverage depends only on the closest available facility (Berman, Drezner and Krass, 2010b).

To overcome the first limitation, gradual coverage models were introduced. Here, coverage decays from 1 to 0 between a minimum and maximum threshold coverage radius. As Drezner, Wesolowsky and Drezner (2004) pointed out, this is suited for EMS systems, where patient outcomes decay with the coverage radius. Multiple coverage functions have been proposed: stepwise functions (Church and Roberts, 1983; Berman and Krass, 2002), linear functions (Berman and Wang, 2011) and S-shaped functions (Drezner, Drezner and Goldstein, 2010). This approach is flexible and can be used to study several objectives, such as minimizing the maximum regret (Berman and Wang, 2011) or maximizing the probability of covered population being above a threshold (Berman, Krass and Wang, 2011). Yet, some authors prefer to use the coverage probability, defined as the probability that a demand node is covered by a facility, to overcome this limitation (Van Den Berg, Kommer and Zuzáková, 2016).

Overcoming the second limitation requires treating the coverage radius as an endogenous parameter: that is, the decision maker can control the coverage radius of a facility that is built (Berman *et al.*, 2009; Davaria *et al.*, 2010). Finally, cooperation is achieved when multiple facilities contribute to coverage by sending a "signal" that decays with distance. A demand node receives multiple signals which are aggregated (e.g. by summation) to provide the overall service (Berman, Drezner and Krass, 2010a).

Thus, the original notion of coverage presents several limitations. In particular, the first and third limitations are particularly important in the EMS context. Unfortunately, the coverage radius of vehicles is not under control of the decision-maker in an EMS system.

4.8.2 PATIENT SURVIVAL

The underlying belief behind coverage models is that quick response times can improve the medical outcomes of patients. Thus, coverage is a proxy for the real goal of EMS: maximize patient survival (Bélanger, Ruiz and Soriano, 2018). Recently, researchers have modelled this objective more explicitly.

Erkut, Erdogan and Ingolfsson (2008) developed a model maximizing the expected number of cardiac arrest survivors. Their model is based on a survival function, describing the survival probability as a function of response time. Comparison with the MCLP and P-Median models provides proof that maximizing survival is more adequate. Later, Knight, Harper and Smith (2012) extended this model to include patients with several medical conditions and, hence, survival functions, while Leknes *et al.* (2017) mixed cardiac arrest survival with coverage, depending on the call's priority. Recently, a simulation study suggested that survival models are superior on coverage and survival metrics when compared to traditional models (Zaffar *et al.*, 2016). However, McLay and Mayorga (2010) concluded that adequately selecting the coverage radius can result in optimal solutions from a survival perspective.

Although survival is a direct description of EMS goals, survival functions may be difficult to obtain. Additionally, maximizing survival can conflict with another goal: equity (Felder and Brinkmann, 2002).

4.8.3 EQUITY

The public expects EMS to provide fair service. Yet, this issue had not received attention until recently (Aringhieri *et al.*, 2017), mainly because there is not yet a consensual measure of equity, and using equity measures alone may lead to inefficient solutions (Smith, Harper and Potts, 2013). However, some authors have acknowledged the importance of equity and formulated different equity objectives.

McLay and Mayorga (2010) provided equity by maximizing coverage in rural areas. Smith, Harper and Potts (2013) combined efficiency and equity, minimizing weighted deviations bellow and above the desired service standard for each region. Chanta, Mayorga and McLay (2014a) studied three equity objectives: minimize the maximum distance from uncovered nodes to their closest facility (P-centre model), minimize uncovered rural areas and minimize total uncovered areas, while Chanta, Mayorga and McLay (2014b) measured equity comparing the survival of patients in different regions.

The Gini-Coefficient, a classical equity measure in Economic and Social systems, has also been used. Drezner, Drezner and Guyse (2009) minimized the Gini-Coefficient based on the Lorenz Curve, where the “good” being distributed is the distance to the closest facility. Equity is achieved when this distance is the same for all nodes. It is concluded that an equitable system may lead to poor overall service. Later, Toro-Díaz *et al.* (2015) considered two equity perspectives: customers and workers. Customer equity was measured using the Gini-Coefficient and the Squared Coefficient of Variation of response times. The Squared Coefficient of Variation of server workload was used to measure server equity.

Another equity measure is the concept of envy, introduced by Chanta *et al.* (2011) in the Minimum P-Envy Location Problem. Envy is a function of the node’s distance to its closest and back-up ambulances when compared to other nodes. Each node feels envy regarding any other demand node, at any level of coverage (primary, back-up, …). The model locates *p* ambulances to minimize total envy. Given the model’s complexity, a Tabu Search is developed. Later, Chanta, Mayorga and McLay (2014b) modified this model to use a survival function instead of distance as the driver of envy.

It is also possible to combine several types of equity. For instance, Cardoso *et al.* (2015) considered four equity areas: equity of access, equity of utilization, socio-economic equity and geographical equity.

Hence, it is possible to conclude that several equity measures have been proposed and that, in absence of a generally accepted metric, this objective must be adapted to the decision-maker’s specific opinion.

4.8.4 MULTI-OBJECTIVE APPROACHES

Recognizing that EMS planners may wish to consider multiple objectives, Multi-Objective models have also been proposed. Sets of conflicting objectives which have been studied include: system costs and demand satisfaction (Harewood, 2002; Zhang and Jiang, 2014); efficiency and equity (Smith, Harper and Potts, 2013; Chanta, Mayorga and McLay, 2014a); primary and back-up coverage (Daskin and Stern, 1981; Storbeck, 1982); population and property coverage (Schilling *et al.*, 1980); population coverage and server back-up coverage (Revelle, Schweitzer and Snyder, 1996); expected coverage and spare server workload (Alsalloum and Rand, 2006).

When dealing with conflicting objectives, there is not an optimal solution. Therefore, several techniques can be applied to study trade-offs between objectives. Some studies aggregate objectives into a single measure using weights (Daskin and Stern, 1981; Hogan and ReVelle, 1986; Zhang and Jiang, 2014), while others derive the Pareto-optimal frontier by varying weights or using the ε-constraint method (Schilling *et al.*, 1980; Revelle and Snyder, 1995; Chanta, Mayorga and McLay, 2014a). A Fuzzy Multi-Objective approach, based on maximising the achievement level of constraints concerning different objectives, can also be used (Tzeng and Chen, 1999; Yang, Jones and Yang, 2007). Finally, Goal Programming is another option: targets are set for each objective and deviations from these targets are minimized (Alsalloum and Rand, 2006; Kanoun, Chabchoub and Aouni, 2010).

4.9 JOINT STRATEGIC, TACTICAL AND OPERATIONAL DECISIONS

Although the distinction between strategic, tactical and operational issues is common, all planning decisions are interrelated and can be made simultaneously (Sung and Lee, 2018). Some authors have attempted to combine them in a single model. One option is to include the dispatching policy in location models by using a fixed preference list (Goldberg and Paz, 1991; Borras and Pastor, 2002; Ingolfsson, Budge and Erkut, 2008; Davoudpour, Mortaz and Hosseinijou, 2014). A preference list is an ordered list of vehicles to serve a demand node, where the first available vehicle on the list will be dispatched. Alternatively, primary and back-up bases can be assigned to each area (Leknes *et al.*, 2017).

Another option is to employ two models, one to capture the dispatching policy and another for the location decision. For instance, Chong, Henderson and Lewis (2016) used a Markovian Decision Process to model dispatching, which is to calculate busy fractions for an optimization model. Toro-Díaz *et al.* (2013), Grannan, Bastian and McLay (2014) and Ansari, McLay and Mayorga (2015) modelled dispatching as fixed preference lists, and use a queueing model to calculate busy and dispatching fractions. Later, Toro-Díaz *et al.* (2015) studied co-location of vehicles and developed a Tabu Search using the model of Budge, Ingolfsson and Erkut (2009) to calculate the queuing aspects of the model. Iannoni, Morabito and Saydam (2009, 2011) developed a Hybrid GA and two Greedy Heuristics to locate ambulances and determine their districts, using an AHQM to capture the dispatching policy. A recent approach is Stochastic Programming. For instance, Sung and Lee (2018) developed a two-stage model in which the location decision is “here-and-now” and dispatching is a “recourse” decision.

It can be concluded that integrating multiple decision-levels has gained traction in the literature, although evidence of the advantage of this approach is still missing.

4.10 OTHER ISSUES FOR EMS VEHICLE LOCATION PLANNING

Besides modelling and developing solution methods, EMS location models often require additional tasks, such as forecasting, demand aggregation and, potentially, the use of simulation, which are briefly discussed next.

Demand and service time forecasting provide inputs to optimization models. Unfortunately, long-term demand forecasting has received little attention, and most studies use historical data or simple averages (Goldberg, 2004). Still, some forecasting approaches have been used, such as statistical inference (Hall, 1971); regression models, using variables as age (McConnel and Wilson, 1998) or demographic, social or geographic factors (Kvalseth and Deems, 1979); time-series models with trends (Channouf *et al.*, 2007) and Winters Exponential Smoothing (Baker and Fitzpatrick, 1986); Single Spectrum Analysis (Vile *et al.*, 2012); and more complex methods, as neural networks (Setzler, Saydam and Park, 2009) or space-time marked point processes (Micheletti *et al.*, 2010).

Forecasting travel times is also important. Kolesar, Walker and Hausner (1975) proposed equations describing mean travel time as a function of distance, concluding that daily fluctuations have a minor impact in travel time. Later, Budge, Ingolfsson and Zerom (2010) proved the validity of these equations for travel time median estimation and use them to develop a model for the probability distribution of travel times, which has been employed by other authors (Alanis, Ingolfsson and Kolfal, 2013).

Another important issue is demand aggregation, which is used to reduce the computational burden of real-sized instances, but can introduce errors (Aringhieri *et al.*, 2017). For instance, Goodchild (1979) showed the significance of this error via simulation. Francis *et al.* (2009) developed a framework for aggregation methods in different classes of location problems and study different error measures.

Finally, the role of simulation is important. Simulation can be used in simulation-optimization, providing estimates of the objective function (Fu, Glover and April, 2005). Here, simulation is an alternative to model stochasticity: a deterministic optimization model suggests vehicle sites, which are refined by simulation in a iterative algorithm (Aringhieri, Carello and Morale, 2016). Alternatively, simulation can be embedded in GA to evaluate the solution's fitness (Zhen *et al.*, 2014). Additionally, simulation is a common tool for model validation (Goldberg *et al.*, 1990). Validation is essential, since modelling requires simplifying assumptions which must be checked. Here, the performance predicted by simulation is usually compared to that of analytical models. Discrete-Event Simulation is more popular, though Agent-Based Simulation can be used as well (Aringhieri, Carello and Morale, 2016). A major difficulty is that simulation requires probability distributions of external parameters (Current and Storbeck, 1988; Pirkul and Schilling, 1991; Shiah and Chen, 2007; Schmid and Doerner, 2010; Yin and Mu, 2012). Alternatively, a set of parameter realizations can be used as scenarios (Aringhieri, Carello and Morale, 2016). An important advantage of simulation regards easy data visualization, thus promoting acceptance by the decision-maker (Henderson and Mason, 2006).

It can be concluded that the impact of these additional issues is paramount to obtain realistic solutions, and care should be taken when performing both preparatory and validation steps.

4.11 CHAPTER CONCLUSIONS

Vehicle location decisions arise at all levels of EMS planning: strategic, tactical and operational. Given their importance in the effectiveness of EMS systems in saving lives, it comes as no surprise that a variety of models have been proposed to address this issue over the last 50 years.

In this chapter, the EMS vehicle location literature is reviewed. It is possible to conclude that a variety of methods exist to model different components of EMS systems, some of which closely resemble the SIEM. Furthermore, the complexity of these models has led to a vast array of solution techniques – exact procedures, heuristics and metaheuristics. The next chapter presents the proposed optimization model, leveraging the approaches presenter in this chapter with the case-study features of chapter 3.

5. MODEL FORMULATION

The goal of this chapter is to propose a Multi-Objective Dynamic MIP Model to address vehicle location planning of EMS systems and, in particular, at the SIEM. This model should leverage the previous knowledge regarding the case-study with the wide range of modelling methodologies which have been covered in the literature review. Additionally, a solution technique is proposed.

Section 5.1 restates the problem statement. Section 5.2 presents the mathematical formulation of the model. After establishing the required notation and sets, subsets, indexed sets and parameters, the objective functions and constraints are presented and explained. Section 5.3 proposes a solution methodology considering one of the model's objectives. Section 5.4 presents the chapter's conclusions.

5.1 PROBLEM STATEMENT

5.1.1 PROBLEM FEATURES AND ASSUMPTIONS

Before presenting the proposed model, it is important to restate the general problem to be addressed. The problem consists of locating EMS stations on a region and assigning emergency vehicles to these stations. Conversely, areas of responsibility of each station/vehicle pair should also be outlined.

The goal is to formulate a strategic plan to allow EMS planners to make decisions by considering both present and future consequences, as well as the changing needs of the population. This enables planners to have a view on how the existing system should evolve. This is also important because choices about deployment of emergency resources are related with many other tasks (e.g. staff planning, financial planning) which must be planned in advance.

It is assumed that the region of interest can be represented by a graph $G = (N, E)$, with nodes (N) and edges (E) connecting them. Customers are grouped into demand nodes, denoted by $d \in D$, which generate emergency requests. It is presumed that these demand nodes have been defined a priori. Depending on the emergency's severity, a given priority, denoted by $p \in P$, is assigned during triage.

Sites for emergency stations are denoted by $f \in F$. Stations can be *existing* (in place at the beginning of the planning horizon), denoted by $f \in F^{exi}$, or potential new stations, denoted by $f \in F^{new}$. Additionally, some stations (existing or potential new) can be changed throughout the planning horizon – these are called *selectable*, denoted by $f \in F^{sel}$ – while others cannot – *non-selectable* stations.

It is also assumed that different emergency vehicles, denoted by $v \in V$, are available. These vehicles are housed at emergency stations. Multiple vehicles of different types can be located at the same station (there is co-location of vehicles). It is assumed that some types of vehicles can be relocated – *selectable* vehicles, $v \in V^{sel}$, while others cannot. A vehicle type may refer not only to the equipment itself, but also to the entity operating it. Therefore, for instance, if a fire department and EMS operate similar vehicles, they can be treated as two separate vehicle types, if such a distinction is required. This can be necessary if they can only be located at different stations, or if the emergencies they can treat are different. Therefore, the model allows considering different entities, with corresponding costs, resource and staff limitations. This can be useful when central planners need to manage both public and private EMS systems or, at least, consider existing private EMS systems in their decisions.

Emergency vehicles are dispatched to emergencies at demand nodes. It is explicitly considered that different emergency vehicles are capable of providing different care levels, $l \in L$. It is assumed that, according to the priority $p \in P$ of an emergency, different care levels, $l \in L^p$, must be provided. For instance, a given priority p may require two BLS vehicles, in which case two care levels are considered, $L^p = \{BLS1, BLS2\}$, representing, respectively, the first and second layers of BLS. Another priority may require one BLS and one ALS vehicle, thus $L^p = \{BLS1, ALS1\}$. Additionally, a vehicle v may be capable of providing more than one care level, $l \in L^v$. For example, a BLS ambulance may be capable of providing all layers of BLS, so $L^p = \{BLS1, BLS2, BLS3, \dots\}$.

Using Queueing Theory terminology, vehicles v at station f can be conceptualized as servers, which are identified by the combination (f, v). Customers, on the other hand, are emergency requests, which "arrive" to these servers after the dispatching operations. Emergency requests result from an emergency originating on demand node d, with priority p, requiring a vehicle to provide care level l. As such, any emergency request can be identified by the combination (d, p, l).

Furthermore, several planning periods, $t \in T$, are considered. These can be years or months, according to the specific planning purposes. There is an existing EMS system in operation at the beginning of the planning horizon. The model will inform planners on how to gradually reconfigure this system over the planning horizon to provide the best possible service. This is important because, in most cases, an existing EMS system is already operating, and it would not be realistic to reconfigure the system in one year. Additionally, this allows predictable fluctuations of demand or travel time (e.g. growing urban areas, road improvements) to be considered for current and future decisions.

Furthermore, time-dependency regarding periods of the day is also included – in other words, working shifts, $s \in S$, are considered. This is because (1) the performance of EMS systems can vary greatly throughout the day and (2) staff is required to operate emergency vehicles, usually working 8-hour shifts. In less active shifts, it may be possible to reduce operating staff without compromising service. For instance, shifts at INEM include morning (00 AM – 08 AM), afternoon (08 AM – 16 PM) and night (16 PM – 24 PM) shifts.

Operating vehicles during these shifts can have different costs, due to crew compensations, and assigning calls during these shifts can also have different costs.

The decisions to be made include the location of stations (binary variables y_f^t), the allocation of available emergency vehicles among these stations (integer variables x_{fv}^t and sh_{fv}^{ts}) and the definition of areas of responsibility of each vehicle/station pair (continuous variables a_{dplfv}^{ts}).

The capacity of these facilities (i.e. number of vehicles they can house) is assumed to be given in advance and not a decision variable. This approach is suited when EMS planners use existing facilities to house vehicles, as is the case of the SIEM. Alternatively, this assumption could be lifted by adding an integer station capacity decision variable to the model. Similarly, the number of vehicles of each type available at each period is assumed to be known, although the number of vehicles to acquire could be added as a decision variable and thus include fleet sizing into the model, subject to budget constraints.

The goal is to provide the best possible service to the population, while maintaining an equitable system with the lowest costs. Since these goals are conflicting, a multi-objective model is appropriate. In particular, three objectives have been included, combining classical and emergent goals aligned with the concerns expressed by INEM practitioners during interviews:

- Demand coverage: maximizing the demand which is successfully serviced by the EMS;
- Cost: minimizing start-up, capacity, operating, relocation and assignment costs;
- Equity: seeking to provide fair service to the population.

5.1.2 RELATIONSHIP TO THE LITERATURE

The model's aim is to combine several aspects of EMS operation into a single framework, flexible enough to accommodate different applications without extensive changes. Many approaches focus only on a particular component of EMS systems, and the literature is still far from a comprehensive and complete model (Beraldi and Bruni, 2009). The proposed model should be viewed as an initial step towards this goal.

Additionally, as mentioned by Church (2002), "(…) there appears to be a lack of handling temporal issues (…). Planning over a number of finite time periods has been the subject of few articles but is more of the norm in actual public application. (…) Most algorithms and heuristics have been tested under the assumption that all facilities are new, i.e., a so-called 'green-field case'. Few heuristics have been designed or tested to optimally add to or modify an existing system, i.e. a 'brown-field case'". Although this gap has been identified over 17 years ago for the general field of location decisions, it remains largely unexplored in the EMS context.

The main contributions of the proposed model can therefore be summarized as:

1. The formulation of a tri-objective model capturing the main concerns – coverage, equity and cost – that drive EMS location decisions, allowing the study of trade-offs between emergent and classical goals;
2. Considering flexibly the possibility of having several types of emergency requests and vehicles which may be dispatched together to a call (multi-dispatch). Most multi-vehicle models in the literature are "hard-coded" to fit a particular application, thus being less flexible.
3. Allowing strategic and tactical relocations (as opposed to operational) by combining micro (shifts) and macro (planning periods) time-dependency. This approach has been applied in other fields, including health-care (e.g. Cardoso *et al.* (2015)), but seldom in EMS literature;
4. Considering an existing EMS system already in operation at the beginning of the planning horizon. Most approaches in the literature assume a "green-field" scenario. Although these models provide information about the ideal vehicle configuration, they do not provide any insight on how the system should progress towards the prescribed solution.

Furthermore, the following features are important:

- Vehicle unavailability is accounted for by using the concept of busy fraction as an endogenous parameter, calculated from the location of stations and vehicles, and limiting the allocation of customers to them accordingly, as in previous studies (Goldberg and Paz, 1991; Shariat-Mohaymany *et al.*, 2012; Leknes *et al.*, 2017).
- The HQM is not used. Since each EMS system is different, a new HQM needs to be developed to fit each particular application (Chiyoshi, Iannoni and Morabito, 2011). This goes against the

desire to have a flexible model. Additionally, such models are complex and usually become of little practical application (Beraldi and Bruni, 2009). In particular, a multi-dispatch, partial back-up HQM with non-homogeneous and co-located servers would be required to model the SIEM. This would be a challenging HQM given the size of the state space.

- Instead of all-or-nothing coverage, the probability that a vehicle at a station can provide coverage to a demand node is used, as in previous studies (Daskin, 1987; Goldberg and Paz, 1991). Thus, the effects of vehicle proximity in patient outcomes are accounted for more explicitly, together with travel time variability. Also, these probabilities can depend on the type of emergency request and care level being provided, because quick response may be more crucial for certain care levels (e.g. speedy response by an ALS first responder might be more important than for a BLS transportation unit). This can easily be translated into the classical notion of coverage by using binary variables instead, taking the value 1 if the vehicle can arrive within the given threshold.

5.2 MATHEMATICAL FORMULATION

In this section, the mathematical formulation of the model is presented, in light of the previous considerations. Before presenting the proposed model, the notation is formally introduced, including sets, subsets, indexed sets and parameters. Then, the objective functions are presented and, finally, the constraints are introduced. A more compact formulation can be found in Appendix C.

5.2.1 SET, SUBSETS AND INDEXED SETS

The following sets are defined:

- $s \in S$: working shifts (e.g. $S = \{Morning, Evening, Night\}$).
- $t \in T$: periods in the planning horizon (e.g. $T = \{Month\ 1, Month\ 2, \dots\}$), $|T|$: number of periods in the planning horizon. $T = 0$ corresponds to the beginning of the planning horizon.
- $d \in D$: demand nodes.
- $p \in P$: emergency priorities (e.g. $S = \{Priority\ 1, Priority\ 2, \dots\}$).
- $v \in V$: vehicle types.
- $l \in L$: care levels (e.g. $L = \{ALS, BLS\}$).
- $f \in F$: emergency station locations.

Together with the following subsets:

- $f \in F^{exi}$: existing emergency station locations.
- $f \in F^{new}$: potential new emergency station locations.
- $f \in F^{sel}$: selectable emergency station locations.
- $v \in V^{sel}$: selectable emergency vehicles.

And, finally, the following indexed sets:

- $v \in V^{f}$: vehicles that can be located at station f.
- $v \in V^{l}$: vehicles capable of providing care level l.
- $f \in F^{v}$: stations where vehicles of type v may be located.
- $l \in L^{p}$: care levels l required by a call of priority type p.

5.2.2 PARAMETERS

The model requires several parameters, which represent input data. These are presented in groups.

Cost Parameters

- $OpeningCost_f^t$: cost of opening a station at site f at the beginning of period t.
- $ClosingCost_f^t$: cost of closing a station at site f at the beginning of period t.
- $CapacityCost_{fv}^t$: average cost per vehicle of type v of operating station f during period t.
- $OperatingCost_v^{ts}$: average cost of operating a vehicle of type v during shift s on period t.
- $AssignmentCost_{dpfvl}^{ts}$: average cost per call from demand node d of priority p for providing care level l from a station at site f with a vehicle of type v during period t and shift s.

Demand Parameters

- Dem_{dp}^{ts}: emergency requests from demand node d of priority p during period t and shift s.

Time Parameters

- $ServiceTime_{dpl}^{ts}$: average time required to provide to a call from demand node d, of priority p, care level l during period t on shift s, excluding travel time.
- $TravelTime_{fvd}^{ts}$: average travel time required for a vehicle departing from a station at site f of type v to arrive at demand node d on period t and shift s.
- θ: maximum travel time in the system.

Coverage Parameters

- $\emptyset_{fvdpl}^{ts}$: probability that a vehicle departing from a station at site $f$of type v can cover a call from demand node d of priority p at care level l during planning period t on shift s.
- $\gamma_{vpl} = \begin{cases} 1, & \text{if a vehicle of type } v \text{ can provide to a call of priority } p \text{ care level } l \\ 0, & \text{otherwise.} \end{cases}$
- $\delta_{fd} = \begin{cases} 1, & \text{if a station at site } f \text{ can be assigned to call on demand node } d. \\ 0, & \text{otherwise.} \end{cases}$
- ε: minimum fraction of coverage that must be provided by a vehicle/station pair to be considered as actively cooperating to serve a demand node's emergency requests.

Resource Parameters

- $StationCap_f^t$: maximum number of vehicles that can be housed at a station at site f during planning period t.
- $VeicAva_v^t$: total number of vehicles of type v available during planning period t.
- $MinVeic_{vf}$: minimum number of vehicles of type v that must be located at a station on site f due to legal requirements.
- $InitialVeic_{vf}$: number of vehicles of type v located at a station on site f at the beginning of the planning horizon.
- $MaxVeicShift_v^{ts}$: maximum number of vehicles of type v that are available to operate during planning period t on shift s.
- $ShiftLength^s$: number of time units in shift s.
- $Days^t$: number of days in planning period t;

- $MaxStations^t$: maximum number of stations that can operate on time period t.

Relocation Parameters

- $MaxOpen^t$: maximum number of stations that can be opened during planning period t.
- $MaxClosed^t$: maximum number of stations that can be closed during planning period t.
- τ_f: minimum amount of time that a station at f must remain open if it is opened during the planning horizon.
- i^t: inflation rate on period t.

Expected Coverage Weights

- W_{pl}^1: weight of covering an emergency of priority p with care level l.
- W_t^2: weight of covering demand during time period t.
- W_s^3: weight of covering demand during working shift s.

5.2.3 DECISION VARIABLES

The model includes primary and auxiliary decision variables, which are presented along their domain.

Primary Decision Variables

The primary decision variables are concerned with the location of emergency stations, the number of vehicles that are allocated to each station and the determination of areas of responsibility for each station/vehicle pair.

- $y_f^t = \begin{cases} 1, & \text{if an emergency station is opened on site } f \text{ during planning period } t. \\ 0, & \text{otherwise.} \end{cases}$
- $x_{fv}^t \in \mathbb{N}_\mathbf{0}$**:** number of vehicles assigned to a station at site f of type v during planning period t.
- $sh_{fv}^{ts} \in \mathbb{N}_\mathbf{0}$**:** number of vehicles assigned to a station at site f of type v during planning period t that are active on working shift s.
- $a_{dplfv}^{ts} \in [0;1]$**:** proportion of demand from node d of priority p for care level l allocated to vehicles positioned at a station at site f of type v during planning period t and working shift s.

Notice that variables x_{fv}^t and sh_{fv}^{ts} are defined only for possible combinations of station locations and vehicles (f, v), thus reducing the number of variables in the model and ensuring that invalid vehicle allocations are not allowed. Also, both variables x_{fv}^t and sh_{fv}^{ts} are required because the relocation of emergency vehicles among shifts is not allowed, due to the INEM's requirements. However, this assumption can be easily lifted by replacing variable x_{fv}^t by sh_{fv}^{ts} and making the necessary changes in the objective functions and constraints.

Additionally, the model allocates demand to vehicles through variables a_{dplfv}^{ts}.This approach is inspired by Leknes *et al.* (2017) and Shariat-Mohaymany *et al.* (2012). The difference is that Leknes *et al.* (2017) only consider two responders per demand node, while Shariat-Mohaymany *et al.* (2012) assume that emergency requests are evenly distributed among stations within the coverage radius. Neither considers multiple emergency priorities and vehicles. Note that variables a_{dplfv}^{ts} are not binary. Therefore, vehicles are allowed to share the workload of a demand node and prevent a static assignment of emergencies to vehicles, aligned with the real operation of EMS systems. This factor is usually considered in the EMS location literature by using the HQM and its variants. In these models, the fraction of dispatches (equivalent to variable a_{dplfv}^{ts}) are determined from the fixed-preference dispatching list and ambulance positions. By letting the model decide these variables endogenously, areas of responsibility are implicitly defined (the greater the variable a_{dplfv}^{ts}, the higher the responsibility). This choice is motivated by the dispatching policy of SIEM, which does not strictly obey a fixed-preference list, since INEM vehicles are dispatched more frequently than outside vehicles.

Auxiliary Decision Variables

These variables are used to control the stability of the EMS system and support the calculation of the equity objective.

- $closed_f^t = \begin{cases} 1, & \text{if a station at site } f \text{ is closed at the beginning of planning period } t. \\ 0, & \text{otherwise.} \end{cases}$
- $opened_f^t = \begin{cases} 1, & \text{if a station at site } f \text{ is opened at the beginning of planning period } t. \\ 0, & \text{otherwise.} \end{cases}$
- $w_{dplfv}^{ts} = \begin{cases} 1, & \text{if calls from demand node } d \text{ of type } p \text{ at care level } l \text{ are assigned to vehicles} \\ & \text{of type } v \text{ at site } f \text{ during planning period } t \text{ and shift } s. \\ 0, & \text{otherwise.} \end{cases}$
- $\overline{ATT} \in \mathbb{R}_0^+$: maximum average travel time for the first responding vehicle over the entire region and planning periods.
- $ATT_{dpl}^{ts} \in \mathbb{R}_0^+$: average travel time for calls from demand node d of type p for providing care level l during planning period t and working shift s.
- $MinATT_{dp}^{ts} \in \mathbb{R}_0^+$: average travel time for the first responding vehicle to calls from demand node d of type p during planning period t and working shift s.
- $E_{fv}^{t+}/E_{fv}^{t-} \in \mathbb{N}_\mathbf{0}$**:** number of vehicles added to/removed from facility f of type v at the beginning of time period t;
- $H_v^{t+}/H_v^{t-} \in \mathbb{N}_\mathbf{0}$**:** total number of vehicles of type v added to/removed from the fleet at the beginning of time period t;
- $b_{fv}^t = \begin{cases} 1, & \text{if vehicles of type } v \text{ are added to site } f \text{ at the beginning of planning period } t. \\ 0, & \text{otherwise.} \end{cases}$
- $c_v^t = \begin{cases} 1, & \text{if vehicles of type } v \text{ are added to the fleet at the beginning of planning period } t. \\ 0, & \text{otherwise.} \end{cases}$
- $\beta_{dpl}^{ts} = \begin{cases} 0, & \text{if for calls from node } d \text{ of type } p, \text{ care level } l \text{ has the shortest average response} \\ & \text{time during planning period } t \text{ and shift } s. \\ 1, & \text{otherwise.} \end{cases}$

5.2.4 OBJECTIVE FUNCTIONS

As stated, three conflicting objectives have emerged from the interviews with INEM and the literature analysis: demand coverage (Z_1), cost (Z_2) and equity (Z_3).

Objective 1: Demand Coverage

As mentioned, coverage is a traditional measure of service level in EMS systems and is also employed by INEM. Therefore, the model seeks to maximize the demand covered across all care levels, demand nodes, emergency priorities, time periods and working shifts. The expression is given by equation (1).

$$Z_1 = \max \sum_{p\in P}\sum_{l\in L} W_{pl}^1 \times \sum_{t\in T} W_t^2 \times \sum_{s\in S} W_s^3 \times \sum_{d\in D}\sum_{f\in F}\sum_{v\in V^f} a_{dplfv}^{ts} \times Dem_{dp}^{ts} \times \emptyset_{fvdpl}^{ts} \quad \textbf{(1)}$$

The objective function expressed by (1) maximizes coverage weighted by three parameters – W_{pl}^1, W_t^2 and W_s^3. The expected coverage provided to emergency requests (d, p, l) is the sum of the fractional coverages provided by all the servers (f, v) which have been allocated to those requests (i.e. those for which $a_{dplfv}^{ts} > 0$). This fractional coverage is obtained by multiplying the amount of demand allocated to each server (given by $a_{dplfv}^{ts} \times Dem_{dp}^{ts}$) by the corresponding value of the coverage probability, $\emptyset_{fvdpl}^{ts}$. Notice that different coverage probabilities can be used, according to the emergency's priority p and the care level l to be provided, thus being able to capture the decision-maker's opinions regarding the importance of vehicle proximity depending on the emergency request.

The busy fraction is not included, since a hybrid model is used, mixing expected coverage and reliability. As will be shown, an upper bound on the unavailability of each vehicle is set instead of including it into the objective function. This way, server equity is promoted alongside coverage with a given reliability.

Weights $W1, W2$ and $W3$ are used to control the importance of maximizing coverage along several dimensions, allowing the decision-maker to let the model prioritize certain emergency requests (given that, in the literature, there is no consensus regarding which call priorities to account for in the model) and certain time periods (as Lei, Cheu and Aldouri (2010))

- $W1_{pl}$: gives the importance of providing care level l to an emergency of priority p. For instance, the decision-maker may feel that providing ALS to an emergent call is of the utmost importance ($W1 = 1$) while providing BLS to the same call is half as important ($W1 = 0.5$);
- $W2_t$: related with the importance of providing coverage during the planning period t. In this case, the decision-maker may consider more important to focus on closer time periods, as optimizing for distant time periods may be less important given that the circumstances are likely to change (but should be taken into account). In this case, $W2_1 > W2_2 > W2_3 > \ldots$
- $W3_s$: controls the importance of providing coverage during the working shift s.

Objective 2: Cost

Additionally, the model minimizes the present value of the total system cost. As shown in previous chapters, cost was not traditionally considered in EMS systems, being usually replaced by a proxy (e.g. resources). However, recent approaches have considered cost explicitly. This is even more important in the SIEM, given that there are multiple entities which are compensated for their services in different ways. Since the model is intended to be used at a strategic level, future costs are discounted. If the planning horizon is shorter, inflation can be ignored. This objective is represented by equation (2).

$$\begin{aligned} Z_2 = min \sum_{t\in T\setminus\{0\}} \frac{1}{(1+i^t)^t} \times \left(\sum_{f\in(F^{exi}\cap F^{sel})} Closed_f^t \times ClosingCost_f^t + \sum_{f\in(F^{new}\cap F^{sel})} Opened_f^t \times OpeningCost_f^t \right) \\ + \sum_{t\in T} \frac{1}{(1+i^t)^t} \times \sum_{f\in F}\sum_{v\in V^f} x_{fv}^t \times CapacityCost_{fv}^t + \sum_{s\in S} sh_{fv}^{ts} \times OperatingCost_v^{ts} \\ + \sum_{t\in T} \frac{1}{(1+i^t)^t} \times \sum_{d\in D}\sum_{p\in P}\sum_{l\in L^p}\sum_{s\in S} a_{dplfv}^{ts} \times Dem_{dp}^{ts} \times AssignmentCost_{dplfv}^{ts} \end{aligned} \quad \textbf{(2)}$$

The cost of the system comprises three components: the fixed capital costs of opening and closing stations during the planning horizon with associated capacity cost for each vehicle; the operating cost of vehicles allocated to each station due to staff requirements; and assignment costs for providing assistance to demand calls. The fleet is assumed to be already in place and, as such, vehicle acquisition costs are not accounted for.

Objective 3: Equity

Finally, the third objective seeks to provide equity, preventing some users to be well served while others are poorly attended. To do so, the worst performance of the system is minimized. From the user's perspective, the time of the first arriving vehicle is usually the most critical, since once this vehicle arrives, emergency care can start. The arrival times of subsequent vehicles are important, but less obvious to the customer. Therefore, to ensure equity, the maximum average travel time of the care level that is provided first, across all shifts and time periods, is minimized, which is simply given by:

$$Z_3 = min\ \overline{ATT} \tag{3}$$

5.2.5 CONSTRAINTS

Finally, it is necessary to present the model's constraints. These constraints are presented in groups: resource constraints, initial system constraints, stability constraints, coverage constraints, utilization constraints, and equity constraints.

Resource Constraints

The first group of constraints is related with the availability of emergency vehicles and limitations to their assignment to stations.

$$sh_{fv}^{ts} \leq x_{fv}^{t}, \qquad \forall f \in F, v \in V^f, t \in T, s \in S \tag{4}$$

$$\sum_{f \in F^v} x_{fv}^{t} \leq VeicAva_v^t, \qquad \forall\, v \in V, t \in T \tag{5}$$

$$x_{fv}^{t} \geq MinVeic_{fv}, \qquad \forall\, f \in F, v \in V^f, t \in T \backslash \{0\} \tag{6}$$

$$\sum_{v \in V^f} x_{fv}^{t} \leq StationCap_f^t \times y_f^t, \qquad \forall\, f \in F, t \in T \tag{7}$$

$$\sum_{v \in V^f} x_{fv}^{t} \geq y_f^t, \qquad \forall\, f \in F, t \in T \tag{8}$$

$$\sum_{f \in F} sh_{fv}^{ts} \leq MaxVeicShift_v^{ts}, \qquad \forall\, v \in V, t \in T, s \in S \tag{9}$$

Constraints (4) state that the number of vehicles of type v on station f that are available during shift s and time period t must be smaller than the total number of vehicles of that type allocated to that station. These constraints mean that, once vehicles have been allocated to a given station, they can either be active or inactive during a shift, but they cannot be relocated from one shift to the next.

Constraints (5) limit the total number of vehicles of each type deployed during each period to the number of available vehicles on that period. Constraints (6), on the other hand, impose a minimum number of vehicles to be allocated to certain stations at all times. These constraints are mostly related with legal requirements, which, for instance, require hospital's emergency departments to have certain types of vehicles available. This requirement may not be met by the existing system. Constraints (7) state that, if a station is opened, then the number of vehicles that can be assigned to it is limited by its capacity (e.g. due to space unavailability). Constraints (8) force at least one vehicle to be assigned to each open station. Otherwise, the model could allow stations to remain open without any vehicle assigned to them. Finally, constraints (9) limit the number of vehicles in operation at each shift, given that, due to crew unavailability, it may not be possible to have all vehicles operating in night shifts.

Initial System Constraints

This set of constraints initializes the decision variables at the beginning of the planning horizon ($t = 0$), thus defining the initial state of the EMS system.

$$x_{fv}^{0} = InitialVeic_{fv}\,, \qquad \forall\, f \in F^{exi}, v \in V^{f} \tag{10}$$

$$y_{f}^{0} = 1\,, \qquad \forall\, f \in F^{exi} \tag{11}$$

Constraints (10) define the number of vehicles of each type allocated to existing stations, while constraints (11) define the open stations. The number of vehicles working at each shift is not initialized since it is assumed that these decisions can be changed from the beginning of the planning horizon.

Stability Constraints

This group of constraints controls the relocation of emergency stations and vehicles throughout the planning horizon, ensuring that the EMS does not oscillate excessively.

$$x_{fv}^{t} \geq x_{fv}^{t-1}\,, \qquad \forall\, v \in \overline{V^{sel}}, f \in F^{v}, t \in T\backslash\{0\} \tag{12}$$

$$y_{f}^{t} \geq y_{f}^{t-1}\,, \qquad \forall\, f \in \overline{F^{sel}}, t \in T\backslash\{0\} \tag{13}$$

Constraints (12) and (13) are very similar. They state that, for non-selectable vehicles ($\overline{V^{sel}}$), assigned vehicles cannot be removed (although additional vehicles may be added). Similarly, non-selectable facilities ($\overline{F^{sel}}$), once opened, must remain in operation until the end of the planning horizon. Remaining facilities and vehicles can be relocated, closed or opened from scratch.

$$y_{f}^{t} - y_{f}^{t-1} = opened_{f}^{t} - closed_{f}^{t}\,, \qquad \forall\, f \in F\,, t \in T\backslash\{0\} \tag{14}$$

$$opened_{f}^{t} + closed_{f}^{t} \leq 1\,, \qquad \forall\, f \in F\,, t \in T\backslash\{0\} \tag{15}$$

$$x_{fv}^{t} - x_{fv}^{(t-1)} = E_{fv}^{t+} - E_{fv}^{t-}\,, \qquad \forall\, f \in F, v \in V^{f}, t \in T\backslash\{0\} \tag{16}$$

$$\sum_{f \in F^{v}} \left(x_{fv}^{t} - x_{fv}^{(t-1)}\right) = H_{v}^{t+} - H_{v}^{t-}, \qquad \forall\, v \in V, t \in T\backslash\{0\} \tag{17}$$

Constraints (14) and (15) keep track of which stations are opened and closed at each time period and constraints (16) and (17) keep track of the vehicles added and removed from each station and the overall emergency fleet. Note that there are infinite possibilities for the differences $E_{fv}^{t+} - E_{fv}^{t-}$ and $H_{v}^{t+} - H_{v}^{t-}$. Since these deviations are not minimized in the objective function, it must be ensured that only one of the E_{fv}^{t} and one of the H_{v}^{t} variables are positive. This is accomplished in constraints (18) to (23).

$$E_{fv}^{t+} \leq b_{fv}^{t} \times StationCap_{f}^{t}\,, \qquad \forall\, f \in F\,, v \in V^{f}, t \in T\backslash\{0\} \tag{18}$$

$$E_{fv}^{t-} \leq (1 - b_{fv}^{t}) \times StationCap_{f}^{t-1}\,, \qquad \forall\, f \in F\,, v \in V^{f}, t \in T\backslash\{0,1\} \tag{19}$$

$$E_{fv}^{1-} \leq (1 - b_{fv}^{1}) \times InitialVeic_{fv}\,, \qquad \forall\, f \in F\,, v \in V^{f} \tag{20}$$

$$H_{v}^{t+} \leq c_{v}^{t} \times VeicAva_{v}^{t}\,, \qquad \forall\, v \in V, t \in T\backslash\{0\} \tag{21}$$

$$H_{v}^{t-} \leq (1 - c_{v}^{t}) \times VeicAva_{f}^{t-1}\,, \qquad \forall\, v \in V, t \in T\backslash\{0,1\} \tag{22}$$

$$H_{v}^{1-} \leq (1 - c_{v}^{1}) \times \sum_{f \in F^{v}} InitialVeic_{fv}\,, \qquad \forall\, v \in V \tag{23}$$

In these constraints, instead of using a Big-M approach, bounds related with capacity and vehicle availability parameters are used in order to tighten the formulation. With these variables properly defined, stability constraints on the relocation of resources can be imposed.

$$\sum_{b=t}^{t+\tau_f^t-1} y_f^b \geq opened_f^t \times \tau_f^t, \qquad \forall\, f \in F, t \in T\backslash\{0\} \tag{24}$$

$$\sum_{f\in F} opened_f^t \leq MaxOpen^t, \qquad \forall\, t \in T\backslash\{0\} \tag{25}$$

$$\sum_{f\in F} closed_f^t \leq MaxClosed^t, \qquad \forall\, t \in T\backslash\{0\} \tag{26}$$

$$\sum_{f\in F} y_f^t \leq MaxStations^t, \qquad \forall\, t \in T\backslash\{0\} \tag{27}$$

$$\left(\sum_{f\in F} E_{fv}^{t+}\right) - H_v^{t+} \leq MaxReal_v^t, \qquad \forall\, v \in V, t \in T \tag{28}$$

Constraints (24) state that, once a station f is opened, it must remain open for the following τ_f time periods. However, at any given time period t, if $t+\tau_f$ is greater that the length of the planning horizon, than $t+\tau_f$ must be replaced by $|T|$. Therefore, τ_f^t can be precomputed as $\tau_f^t = \min\,(\tau_f;\ |T|-t+1)$.

For instance, in Figure 12, stations A and B should remain opened for 4 periods once opened. Station B starts operating at the beginning of period 2, and remains open in the following 4 periods, until the end of period 5. Afterwards, it can remain open or be closed. On the other hand, station A only starts operating on period 4 and, since there are only 3 time periods remaining, it will operate until the end.

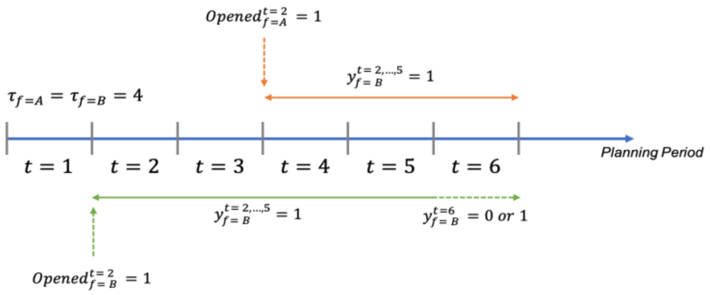


Figure 12 - Example for constraint (24).

In this case, it results that:

$$\tau_{f=A}^{t=4} = \min\left(\tau_f;\ |T|-t+1\right) = \min(4; 6-4+1) = \min(4;3) = 3$$

Constraints (25) and (26) place a limit on the amount of stations that can be opened and closed at each time period, while constraints (27) limit the total number of stations, all according to the decision-maker's wishes. Finally, constraints (28) place a limit on the total number of vehicles of each type that can be relocated during each time period. Not all these constraints need to be implemented simultaneously. The appropriate constraints must be chosen according the decision-maker's wishes.

Coverage Constraints

Coverage constraints deal with the assignment of emergency requests to vehicles at stations, ensuring those assignments are possible.

$$\sum_{f\in F}\sum_{v\in V^f} a_{dplfv}^{ts} \leq 1, \qquad \forall\, d \in D, p \in P, l \in L^p, s \in S, t \in T \tag{29}$$

$$\sum_{l\in L^p} a_{dplfv}^{ts} \leq sh_{fv}^{ts}, \qquad \forall\, d \in D, p \in P, f \in F, v \in V, s \in S, t \in T \tag{30}$$

$$a_{dplfv}^{ts} \leq \gamma_{vpl} \times \delta_{fd}, \qquad \forall \in D, p \in P, l \in L, t\, f \in F, v \in V, s \in S, t \in T \tag{31}$$

$$w_{dplfv}^{ts} \geq a_{dplfv}^{ts}\,, \qquad \forall \in D, p \in P, l \in L, t\, f \in F, v \in V, s \in S, t \in T \tag{32}$$

$$\varepsilon \times w_{dplfv}^{ts} \leq a_{dplfv}^{ts}\,, \qquad \forall \in D, p \in P, l \in L, t\, f \in F, v \in V, s \in S, t \in T \tag{33}$$

Constraints (29) state that the total fraction of (d, p, l) requests assigned to station/vehicle pairs cannot exceed 1, that is, it is not possible to assign more than the existing demand. Constraints (30) prevent any station/vehicle pair (f, v) from being responsible for providing more care levels to the same call than the number of existing vehicles. Suppose that there is only one vehicle in a given station. If a call of type p requires two care levels, ALS and BLS, then it is not possible to have $a_{dp,\mathrm{ALS},fv}^{ts} + a_{dp,\mathrm{BLS},fv}^{ts} = 0.5 + 1 \geq 1$. If that were the case, then the single vehicle would have to provide two different care levels to the same call: BLS care to all calls and ALS to half of them.

Constraints (31) only allow requests (d, p, l) to be assigned to a vehicle of type v if two conditions are met simultaneously: that vehicle type is capable of providing care level l to a call of priority p $(\gamma_{vpl} = 1)$ and vehicles at station f are allowed to respond to calls on demand node d $\left(\delta_{fd} = 1\right)$. The first condition may not be met if the vehicle does not have the necessary medical resources, while the second condition may not be met if, for instance, firefighter operators do not respond to calls outside their area. Constraints (32) assign the proper value to variables w_{dplfv}^{ts} (linking constraints), while constraints (33) force a station to cover a minimum amount of demand, if it cooperates to serve a demand node.

The previous constraints assign emergency requests to station/vehicle pairs. However, an assignment of which guarantees coverage with a given reliability is desirable, as described next.

Utilization Constraints

This group of constraints calculates the expected utilization of vehicles and guarantees that enough vehicles are assigned to emergency requests of a given priority to ensure coverage. The utilization of vehicles of type v on station f during planning period t and working shift s (ρ_{fv}^{ts}) is calculated by dividing the total amount of time the vehicle is expected to be busy serving calls (determined by the assignment variables a_{dplfv}^{ts}) by the total availability time. In doing so, it is assumed that the workload arriving for vehicles of type v on station f is distributed evenly among vehicles of that type at that station.

$$\rho_{fv}^{ts} = \frac{\sum_{d\in D}\sum_{p\in P}\sum_{l\in L} Dem_{dp}^{ts} \times a_{dplfv}^{ts} \times \left(TravelTime_{fvd}^{ts} + ServiceTime_{dpl}^{ts}\right)}{sh_{fv}^{ts} \times ShiftLength^{s}}, \qquad \forall\, f \in F, v \in V^{f}, s \in S, t \in T \tag{34}$$

Having defined ρ_{fv}^{ts}, constraints to guarantee coverage reliability can be developed. An idea similar to Shariat-Mohaymany *et al.* (2012) is used. However, instead of assuming server independence, it is assumed that, from the perspective of an emergency, servers operate as a M/M/N/∞ queueing system, where N is the number of servers assigned to that customer. This implies assuming that requests happen according to a Poisson process, service time is exponentially distributed (which is an important assumption) and that requests will wait if all servers are busy (which is reasonable, since there is no secondary system to which they can be transferred). Consequently, the Erlang-C Formula is applied, which gives the blocking probability (P_B) of the system, that is, the probability that an emergency request finds all of its N assigned vehicles busy, given that these vehicles have a busy fraction equal to ρ:

$$P_B = \frac{\frac{\rho^N N^N}{N!\,(1-\rho)}}{\left[\sum_{n=0}^{N-1} \frac{(N\rho)^n}{n!} + \frac{\rho^N N^N}{N!\,(1-\rho)}\right]} \tag{35}$$

Therefore, if coverage with reliability α is to be ensured, it must be guaranteed that $P_B \leq 1 - \alpha$ by choosing values of ρ^{max} and N according to (35). The parameter N should be the number of vehicles that dispatchers analyse before dispatching a vehicle, thus effectively representing the number of vehicles responsible for that request, and ρ^{max} follows from (35) and the chosen reliability level.

Table 2 - Maximum workload per server for different numbers of servers and reliability levels.

	Reliability (α)				
Servers (N)	**80%**	**85%**	**90%**	**95%**	**99%**
1	0.20000	0.15000	0.10000	0.05000	0.01000
2	0.37015	0.31390	0.25000	0.17110	0.07325
3	0.46433	0.41033	0.34667	0.26253	0.14303
4	0.52550	0.47475	0.41325	0.32975	0.20250
5	0.56940	0.52140	0.46260	0.38100	0.25180

Then, an upper bound on the unavailability probability of each vehicle (constraints (36)) is imposed and a lower bound on the number of vehicles that must share the responsibility of attending to an emergency request (constraints (37)). If required, these values can be different for each shift and time period combinations because, for instance, the decision-maker may wish to provide higher coverage reliability in busier periods.

$$\sum_{d \in D} \sum_{p \in P} \sum_{l \in L} Dem_{dp}^{ts} \times a_{dplfv}^{ts} \times \left(TravelTime_{fvd}^{ts} + ServiceTime_{dpl}^{ts}\right) \leq \rho^{max} \times sh_{fv}^{ts} \times ShiftLength^{s}, \qquad \forall\, f \in F, v \in V^f, s \in S, t \in T \tag{36}$$

$$\sum_{f \in F} \sum_{v \in V^f} \sum_{l \in L^p} w_{dplfv}^{ts} \geq N\,, \qquad \forall\, d \in D, p \in P, s \in S, t \in T \tag{37}$$

Equity Constraints

Finally, this group of constraints calculates the required values for the equity objective function.

$$ATT_{dpl}^{ts} = \left(\sum_{f \in F} \sum_{v \in V^f} a_{dplfv}^{ts} \times TravelTime_{fvd}^{ts}\right) + \left(1 - \sum_{f \in F} \sum_{v \in V^f} a_{dplfv}^{ts}\right) \times \theta, \qquad \forall\, d \in D, p \in P, l \in L^p, s \in S, t \in T \tag{38}$$

$$MinATT_{dp}^{ts} \geq ATT_{dpl}^{ts} - \theta \times \beta_{dpl}^{ts}, \qquad \forall\, d \in D, p \in P, l \in L^p, s \in S, t \in T \tag{39}$$

$$\sum_{l \in L^p} \beta_{dpl}^{ts} = |L^p| - 1\,, \qquad \forall\, d \in D, p \in P, l \in L^p, s \in S, t \in T \tag{40}$$

$$\overline{ATT} \geq MinATT_{dp}^{ts}\,, \qquad \forall\, d \in D, p \in P, s \in S, t \in T \tag{41}$$

Constraints (38) calculate the average travel time for each emergency request at each time period. Similarly to Cardoso *et al.* (2015), unassigned requests (i.e. those not assigned to any station/vehicle pair) are penalized to a maximum value of travel time. Constraints (39) and (40) calculate, for each call priority of each region, the minimum average travel time across all care levels (that is, the fastest responder). Finally, constraints (41) assign the largest of these travel times to $\overline{ATT}$.

5.3 SOLUTION APPROACH

Given the model's complexity and expected size of real instances, traditional solvers, such as CPLEX, employing optimal solution methods may consume significant time. Even though the model is intended for strategic/tactical planning, which is performed less frequently and therefore where computational speed is less crucial, it is still important to seek alternative solution procedures which streamline model solution. For instance, this is necessary if the model is to be integrated in a decision-support system.

As shown in the previous chapter, multiple heuristic approaches have been developed to address EMS location problems. These heuristics seek to explore underlying structure of the model. In this section, a hybrid heuristic which iterates between two sub-problems resulting from decomposing the original model is proposed to solve the model for the expected coverage (Z1) objective.

5.3.1 MODEL DECOMPOSITION

Note that the model integrates three decisions: station selection, allocation of vehicles to stations and assignments of nodes to vehicles. Although taking these decisions simultaneously is expected to yield better solutions, they can be made sequentially. In particular, stations can be selected, and vehicles allocated to these sites before assigning demand. This observation motivates the proposed heuristic, which consists of decomposing the original model in two sub-models and iterate between them.

The first sub-model (SP1) decides the location of facilities and allocation of vehicles. This model is, in fact, a traditional gradual coverage model with multiple customer and server types. For this purpose, instead of maximizing expected coverage (which takes into account demand assignment variables), potential expected coverage is maximized. This is achieved by replacing the variables a_{dplfv}^{ts} with sh_{fv}^{ts} in the expression of the objective function. Furthermore, all constraints related with demand assignment are lifted. Accordingly, the formulation of this model is as follows:

SP1:

$$Z_1 = \max \sum_{p \in P} \sum_{l \in L} W_{pl}^1 \times \sum_{t \in T} W_t^2 \times \sum_{s \in S} W_s^3 \times \sum_{d \in D} \sum_{f \in F} \sum_{v \in V^f} sh_{fv}^{ts} \times Dem_{dp}^{ts} \times \emptyset_{fvdpl}^{ts}$$

Subject to constraints (4) to (28).

Once the sh_{fv}^{ts} variables are fixed, they can be inputted into a second sub-problem, which assigns demand to these vehicles without exceeding vehicle utilization constraints, considering the remaining constraints and the original objective function. The second sub-model (SP2) is as follows:

SP2:

$$Z_1 = \max \sum_{p \in P} \sum_{l \in L} W_{pl}^1 \times \sum_{t \in T} W_t^2 \times \sum_{s \in S} W_s^3 \times \sum_{d \in D} \sum_{f \in F} \sum_{v \in V^f} a_{dplfv}^{ts} \times Dem_{dp}^{ts} \times \emptyset_{fvdpl}^{ts}$$

Subject to constraints (29) to (33) and (36) to (37).

5.3.2 PROPOSED HEURISTIC

Instead of solving SP1 and SP2 sequentially, one can iterate between these two models, adding constraints to SP1 which account for the results of inputting the sh_{fv}^{ts} of the previous iteration into SP2, approximating this solution to the optimal solution. For this purpose, note that the optimal solutions of SP1 consider nodes covered multiple times similarly to multiple nodes covered once. As such, this model tends to add more vehicles in regions where the demand is high, regardless of the number of vehicles already positioned at those sites. However, after a certain point, adding more vehicles to a region may not provide any actual benefit in the original model, since the existing vehicles are sufficient to cover all requests reliably. Therefore, adding constraints to the SP1 which prevent this situation should lead to better, more distributed solutions. The heuristic procedure is outlined in Table 3.

Table 3 - Pseudo-code of the proposed two-stage hybrid heuristic.

Algorithm Hybrid Heuristic

Initialize: $BFS \leftarrow 0,\ iter \leftarrow 1, cons \leftarrow 0$
Set $N,\ \varepsilon,$
While $(iter \leq Maximum\ iterations)$ ***and*** $(cons \leq Maximum\ iterations\ without\ improvement)$ ***do:***
    Solve $SP1$ *and retrieve optimal decision-variable* sh^{iter}
    Solve $SP2$ *and retrieve optimal solution value* Z^*
    If $(Z^* \geq BFS))$ ***then:***
        Set $BFS \leftarrow Z^*$
        Set $cons \leftarrow 0$
    Else:
        Set $cons \leftarrow cons + 1$
    For Each $(sh_{fv}^{ts\,iter})$ ***do:***
        Compute: $rank_{fv}^{ts\,iter} \leftarrow \sum_{p\in P}\sum_{l\in L} W_{pl}^1 \times \sum_{t\in T} W_t^2 \times \sum_{s\in S} W_s^3 \times \sum_{d\in D}\sum_{f\in F}\sum_{v\in Vf} sh_{fv}^{ts} \times Dem_{dp}^{ts} \times \emptyset_{fvdpl}^{ts} - \sum_{p\in P}\sum_{l\in L} W_{pl}^1 \times \sum_{t\in T} W_t^2 \times \sum_{s\in S} W_s^3 \times \sum_{d\in D}\sum_{f\in F}\sum_{v\in Vf} a_{dplfv}^{ts} \times Dem_{dp}^{ts} \times \emptyset_{fvdpl}^{ts}$
    Sort $sh_{fv}^{ts\,iter}$ according to $rank_{fv}^{ts\,iter}$
    For *the first* N $sh_{fv}^{ts\,iter}$ ***do:***
        Add *constraint* $sh_{fv}^{ts} \leq \max(1,\ sh_{fv}^{ts\,iter} - 1)$
    Update $N \leftarrow round\left(\frac{N}{\varepsilon}\right)$
    Set $iter \leftarrow iter + 1$

The proposed heuristic starts by solving SP1 and SP2 in sequence. Comparing both solutions, the heuristic accesses which station/vehicle pairs have been most overestimated in the objective function of SP1 by comparing the terms of both objective functions. For the N most overestimated pairs, a cut is added to SP1 in order to promote more diluted solutions. Furthermore, the value of N is updated in order to improve the speed of the heuristic. The key insight is that, in the beginning of the optimization, many station/vehicle pairs are overestimated by SP1. However, as cuts are introduced, the solution starts converging towards the optimal solution, and smaller adjustments are required. Therefore, the number of constraints added to SP1 decays as the heuristic progresses.

5.4 CHAPTER CONCLUSIONSS

In this chapter, the problem is restated, and a Multi-Objective Dynamic Mixed-Integer Programming model is developed. This model is proposed to aid vehicle planning by considering three levels of decision: selection of stations, assignment of vehicles to stations and allocation of demand to vehicle/station pairs. Therefore, it encompasses both strategic and tactical decisions. Furthermore, three objectives – coverage, cost and equity – are included, alongside different constraints capturing restrictions to the desired plan. Different vehicles and call priorities, as well as multiple time periods and shifts, are accounted for, thus allowing for the gradual reconfiguration of the existing system. A two-stage hybrid heuristic procedure is also proposed, seeking to iterate between two subproblems of the original formulation to derive good feasible solutions considering only the first objective (coverage).

The following chapter describes the data collection and treatment procedures required to apply the proposed model to the case study.

6. DATA COLLECTION AND ANALYSIS

This chapter introduces the data collection and analysis procedures required to apply the model to the case study. Since not all necessary data could be made available by INEM, section 6.1 lists all the assumptions required to estimate model inputs from available data. Section 6.2 describes data collection procedures, while section 6.3 describes the required methods to transform the data into parameters. Finally, section 6.4 presents the chapter's conclusions.

6.1 ASSUMPTIONS AND LIMITATIONS

This section summarizes the assumptions required to estimate model inputs from the available real-world data. Firstly, due to privacy concerns, historical records identify each emergency's location by its postal code. Although it is possible to geo-reference each postal code, this approximation may introduce errors. Moreover, for some emergencies, only a shorter (four digit) postal code is available, corresponding to larger regions within each city. Since the number of entries in the records without the complete postal code is significant (around 40%), these cannot be discarded. Given that the number 7-digit postal code emergencies is significant in all regions – above 60% – it is assumed that 4-digit emergencies are distributed similarly to 7-digit emergencies of the same region. This assumption is possible because there is no pattern regarding which areas are identified with a 4-digit code.

Another limitation is that INEM cannot provide travel time data. With this information, it would be possible to extrapolate travel times to feed the model, compute different coverage probabilities and derive indicators to assess the quality of the current system and compare it to the proposed solution. To overcome this limitation, Google Maps is used to calculate the distance between the emergency stations and the centroid of each demand region, and an empirical model for emergency vehicle travel time (Budge, Ingolfsson and Zerom, 2010) is applied to compute the travel time distribution. Furthermore, it is also assumed that:

- Euclidean distances can be used to approximate travel times for demand aggregation purposes, given that the road networks are dense in urban areas, such as Lisbon and Setúbal;
- Emergency vehicles always use the shortest-time route;
- Since all considered vehicles have 4 wheels, travel times and coverage probabilities are independent of the vehicle type;
- Given that Lisbon and Setúbal are not expected to undergo structural changes in their road infrastructures, travel times are also assumed to be independent of the month of the year;
- Station capacities are also presumed to remain equal throughout the planning horizon;

Finally, cost information is not available since the GPCG is not responsible for cost analysis at INEM. As such, costs are estimated from public reports. It is assumed that costs remain constant throughout the planning horizon, that opening and closing stations bears no cost and that NINEM costs are similar to RES. Furthermore, average fuel and medical supplies consumptions are used to estimate the assignment costs of INEM vehicles. The potential impact of uncertainty regarding these parameters is studied in Chapter 7.

6.2 DATA COLLECTION PROCEDURES

To apply the proposed model to INEM's case-study, relevant data needs to be collected and treated to generate model inputs. Employed data collection procedures include:

1. **Interviews with INEM practitioner**: during six interviews with a planning technician of the GPCG, the SIEM is characterized and several inputs are defined, such as resource availability, coverage targets, objective function weights and system reconfiguration parameters;
2. **Analysis of historical emergency request records**: historical records from 2017 and half 2018 (the most recent available records) are collected and analysed to develop emergency request forecasts for the planning period;
3. **Public platforms**: such as Google Maps, are used to estimate parameters for which historical data is not available.

6.3 PLANNING PERIOD AND REGION CHARACTERIZATION

In line with the current practice, and in order to support the formulation of INEM's yearly vehicle location plan, a 12-month planning horizon is analysed (the year of 2020). Additionally, three 8-hour shifts are considered: morning (00:00 AM – 08:00 AM), evening (08:00 AM – 04:00 PM) and night (04:00 PM – 12:00 PM). Therefore, an average day within each month is considered, and three shifts are studied.

Since the model conceptualizes the region of interest as a discrete space, emergency requests must be aggregated in demand areas and station locations must be known in advance. This section characterizes Lisbon and Setúbal, identifying demand areas and station locations. Geographic analysis, including demand aggregation, is conducted using a Geographic Information System (GIS) software called QGIS. Georeferencing is accomplished using the Google Geocoding API in a Python script.

6.3.1 DEMAND AREAS

The demand point aggregation problem consists of aggregating Demand Points (DP) into Aggregate Demand Points (ADP). Besides minimizing the computational cost, demand aggregation eases data collection and modelling, as well as reduces statistical uncertainty. However, demand aggregation introduces errors into the model. Therefore, a trade-off must be considered: using more demand points reduces the aggregation error but creates more difficult models (Francis *et al.*, 2009).

It is important to highlight that, since the formulation of the coverage (Z_1) and cost objectives (Z_2) use an additive structure, aggregation errors may cancel out. However, this is not the case for equity (Z_3).

In order to aggregate demand requests, emergency records from 2017 and half of 2018 are analysed. As mentioned, emergency records identify each emergency's location through its postal code. Portuguese postal codes are composed of 7 digits. The first 4 digits identify a region within a city, while the remaining 3 identify street segments or collections of streets. Figure 13 shows Lisbon's 4-digit and 7-digit postal code areas (Voronoi Polygons[2]). A limitation in the available records is that some emergencies are identified only by the 4-digit postal code. To overcome this limitation, a first possibility would be to use the centroids of the 4-digit postal code areas as ADPs. However, as Figure 13 shows, these areas are excessively large and non-uniform, and would likely introduce large aggregation errors.

[2] Voronoi polygons partition a plane based on points such that the region associated with each point contains all locations closer to that point than any other. In this case, each postal code's georeferenced location is used.

Another option is to discard the 4-digit observations. However, these account for 44.52% of the entries in Lisbon's records, and 37.15% in Setubal's, meaning that these observations are significant and should not be discarded. Consequently, the 4-digit postal codes are converted into 7-digit postal codes by assuming that 4-digit emergencies are distributed similarly to the 7-digit emergencies of the same 4-digit region. A third possibility is to use 7-digit areas as ADPs. However, as Figure 13 also illustrates, there is a significant number of 7-digit postal code areas, which would likely lead to an intractable model.

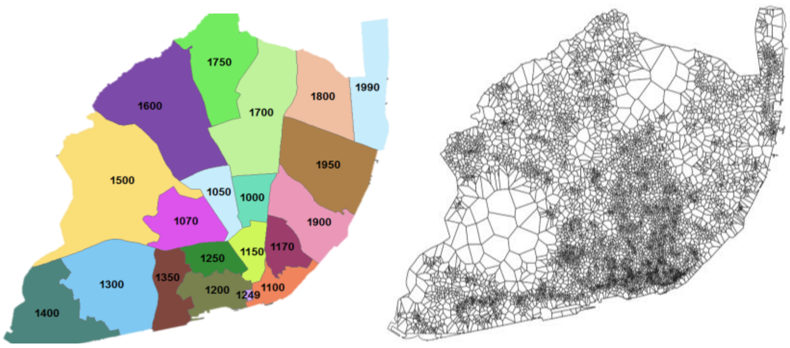


Figure 13 - Postal code areas in Lisbon (Right: 4-digit postal code areas; Left: 7-digit postal code areas).

Therefore, alternative aggregation schemes had to be considered. Unfortunately, the literature has concentrated more on aggregation errors of P-median than covering models (Francis *et al.*, 2009). Nevertheless, the following five aggregation algorithms are identified: K-Means (KM) (MacQueen, 1967), Pick the Farthest (PTF) (Daskin *et al.*, 1989), Independent Projection Algorithm (IPA) (Emir-Farinas and Francis, 2005), Approximate Common Reachability Set (ACRS) method (Jang and Lee, 2015) and Density-Based Spatial Clustering of Applications with Noise (DBSCAN) (Ester *et al.*, 1996).

Jang and Lee (2015) provide a comparison of the IPA, PTS, K-Means and the ACRS for covering problems and conclude that the IPA and PTS present the highest aggregation errors. Therefore, these methods are excluded. On the other hand, the ACRS requires an explicit unique gradual coverage function. Although the proposed model uses a gradual coverage objective, it is based on travel time probability estimation and not on a single explicit gradual coverage function. This excludes the ACRS.

As such, only the DBSCAN and the KM methods are tested. In both cases, Euclidean distances among georeferenced postal codes are used. Since the road network in both areas is dense, it is assumed that the path followed between any two points is approximately linear. In rural areas, this assumption may not hold. Regarding parameter definition for the DBSCAN, the minimum number of points in a cluster is set to 1, so that all DPs are assigned to one ADP. The maximum distance between DPs of one cluster is adjusted to achieve the same number of clusters as the KM. The algorithms are applied using the corresponding plugin in QGIS. The number of ADPs is set to 1%, 2% and 5% of the number of DPs, in order to allow testing the proposed model for different instance sizes. Therefore, three demand aggregation possibilities are developed for each city, as presented in Table 4.

Table 4 - Cluster partition possibilities and corresponding number of aggregate demand points.

Cluster Partition	Lisbon	Setúbal
1%	33	27
2%	66	55
5%	165	137

The results are presented in Figure 14 for 165 (5%) ADPs in Lisbon.

Figure 14 - Demand aggregation areas for Lisbon using the DBSCAN (left) and KM (right) methods.

Due to the high density of postal codes in the centre of Lisbon, the DBSCAN results in very heterogeneous demand regions. In order to achieve smaller regions near the city centre, a short distance between points of the same cluster would have to be used, resulting in a prohibitively large number of ADPs. The same problem is encountered in Setúbal. The KM, on the other hand, provides a more appealing partitioning, with identically sized clusters evenly spread throughout the region. Therefore, given that the results of the DBSCAN are unsuitable, the ADPs resulting from KM are used as demand areas for the case-study regions. The final clusters using KM for 1%, 2% and 5% of the number of DPs are displayed in Figure 15.

Figure 15 - Final cluster partitions and corresponding centroids for Lisbon and Setúbal using KM (Above: 1% (33), 2% (66) and 5% (165) clusters in Lisbon; Bellow: 1% (27), 2% (54) and 5% (135) clusters in Setúbal).

6.3.2 EMERGENCY STATIONS

Station locations can be divided into two sets: existing stations and potential new stations. Existing stations in Lisbon and Setúbal are presented in Figure 16.

Unfortunately, the GPCG does not have a list of potential sites for new stations. Therefore, these stations had to be surveyed. For this purpose, a set of criteria is defined, based on interviews, to establish rules for identifying potential new stations.

Currently, INEM seeks to explore public facilities as stations, mainly due to their increased flexibility and lower costs. Facilities which belong to the Ministry of Health are preferred. Given these preferences, two levels of potential stations are defined:

- **Level 1:** hospitals and primary health centres as well as INEM, firefighter and Red Cross facilities;
- **Level 2:** schools and police stations meeting certain infrastructure conditions.

Using Google Maps to survey the regions of interest and validating with INEM, a set of potential stations is collected and georeferenced using the Google Geocoding API. This resulted in a total of 95 potential new sites for Lisbon and 11 in Setúbal, which are presented in Figure 17.

The capacity of each existing station is available from INEM. On the other hand, the capacity of new stations is estimated as two vehicles per station, except for some stations in which space constraints limit the capacity to only one vehicle.

Besides determining the location of existing and potential stations, the maximum number of stations on each period are set to 20 in Lisbon and 5 in Setúbal, across all months except June – September, in which 23 stations are allowed in Lisbon and 6 in Setúbal. Additionally, the number of stations which can be opened and closed in each month are presented in Table 27 of Appendix E.

Figure 16 - Existing emergency stations in Lisbon (above) and Setúbal (bellow).

Figure 17 - Surveyed potential new emergency stations (Green: level 1; Orange: level 2).

6.4 EMERGENCY VEHICLES

Only vehicles capable of answering P1 and P3 calls are considered (NINEM, RES, PEM, AEM, SIV and VMER). MEMs are excluded because only one of these vehicles operates in Lisbon and none in Setúbal, and their positions have to remain unchanged due to external factors. NINEM, RES and PEM are non-selectable, which means that they cannot be moved once they are assigned to a given region. AEMs, VMERs and SIVs are selectable. The initial system state is characterized by the number of vehicles housed in each station. This is the solution currently in place and is presented in Table 5. Seasonal PEMs, which are positioned during the summer, have not been included because their position is variable.

Table 5 - Current distribution of vehicles among emergency stations.

STATIONS		VMER	SIV	AEM	PEM	NINEM	TOTAL
LISBON		**3**	**1**	**14**	**2**	**6**	**26**
1	*INEM Headquarters*		1				1
2	*GNR Reg. Cavalaria - Ajuda*			1			1
3	*Esquadra PSP - Bairro Boavista*			1			1
4	*Centro Saúde Lóios - Olivais*			3			3
5	*Hospital Curry Cabral*			2			2
6	*Escola S.D. Benfica*			1			1
7	*GNR Brigada Fiscal - Beato*			1			1
8	*GNR Brigada Trânsito - Alcântara*			1			1
9	*Hospital Egas Moniz*			1			1
10	*Hospital São Francisco Xavier*	1					1
11	*Hospital São José*	1					1
12	*Hospital Santa Maria*	1					1
13	*INEM - R. Infante D. Pedro*			3			3
14	*Reg. Sapadores Bomb. - Av D. Carlos I*				2		2
15	*Volunteer Firefighters Ajuda*					1	1
16	*Volunteer Firefighters Beato*					1	1
17	*Volunteer Firefighters Campo De Ourique*					1	1
18	*Volunteer Firefighters Cabo Ruívo*					1	1
19	*Volunteer Firefighters Lisboa*					1	1
20	*Volunteer Firefighters Lisbonenses*					1	1
SETÚBAL		**1**		**2**	**5**		**8**
1	*Hospital de Setúbal*	1					1
2	*Hospital Psiquiátrico Setúbal*			2			2
3	*Volunteer Firefighters Setúbal - Sede*				2		2
4	*Volunteer Firefighters Setúbal - Azeitão*				1		1
5	*Cruz Vermelha Setúbal*				2		2
TOTAL		**4**	**1**	**16**	**7**	**6**	**34**

The availability of emergency vehicles on each month and each shift is presented in Table 28 and Table 29 of Appendix E. Note that seasonal vehicles are introduced in the summer months. Additionally, NINEMs, VMERs and SIVs are available on all shifts, while no RESs are available on any city. AEMs and PEMs are partially available in the evening and night shifts.

Vehicles are capable of providing BLS and ALS. BLS is provided by all vehicles except VMERs. ALS is provided by VMERs and SIVs. Additional information regarding the stations in which each vehicle can be positioned is necessary. For this purpose, stations are classified in 4 categories, as in the legislation: health units, INEM, Firefighters/RC, police/schools. Table 6 describes the possible allocations of vehicles to stations.

Table 6 - Allowed allocations of emergency vehicles to emergency stations.

	Health Units	INEM HQ	Police/Schools	Firefighters/RC
VMER	Yes	Yes	No	No
SIV	Yes	Yes	No	No
AEM	Yes	Yes	Yes	Yes
PEM	No	No	No	Yes
RES	No	No	No	Yes
NINEM	No	No	No	Yes

6.5 EMERGENCY REQUESTS

In order to derive the demand parameters, the most recent information on historical emergency records of Lisbon and Setúbal is used. 115.951 records are available for Lisbon and 18.675 for Setúbal. These records describe, for every emergency, the corresponding date and time, postal code, priority, dispatched vehicles and the time these vehicles became available ("Base Time").

In a preliminary step, duplicate records are removed, and the remaining records are filtered to include P1 and P3 calls. Auxiliary fields are also calculated, including shift, month, total service time (difference between start and base time), time between successive calls (measured at the first dispatch) and the care level for P1 calls. A two-step approach is proposed:

1. **Exploratory analysis**: analysing emergency requests in order to verify previous assumptions regarding the underlying patterns. Data plots and statistical tests are used.
2. **Model fitting and forecasting**: an appropriate model should be chosen and applied to predict the emergency requests over the planning horizon.

The following two sections describe the results of the two-stage approach, which is performed using the statistical software R (R Core Team, 2017). Distribution fitting and statistical tests use the *fitdistrplus* package (Delignette-Muller *et al.*, 2019)

6.5.1 EXPLORATORY ANALYSIS

Firstly, the initial assumption that the demand depends both on the time of the day and month is validated. Since the goal is to forecast call volumes, and in order to remove the effect of variable month length, the analysis focuses on the daily volume of calls. Boxplots of daily volume of calls, grouped by month (example in Figure 18) and shift, as well as the evolution of the arrival rate during the day (Figure 19), suggest that the demand pattern is not stationary with month nor time of the day.

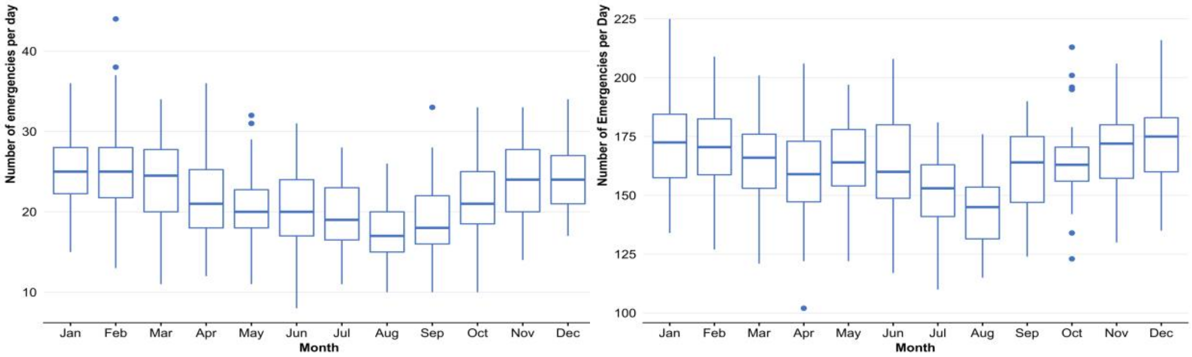


Figure 18 - Boxplots of the daily volume of calls for each month in Lisbon (Left: P1 calls; Right: P3 calls).

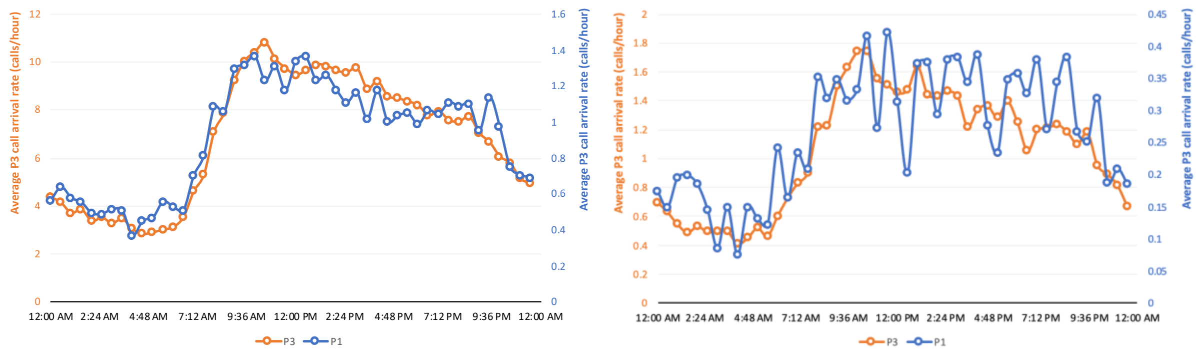


Figure 19 - Evolution of the average call arrival rate during the day for Lisbon (left) and Setúbal (right).

The call arrival rate is usually greater in during the day (approx. 8 AM to 10 PM) and considerably lower during the night periods. Additionally, there tend to be more calls during the winter and less in the summer. This suggests that the call generating process is not stationary.

In order to statistically verify this assumption, data is aggregated by month and shift, the mean and standard deviation (SD) of the daily volume of calls are calculated (Appendix D). Preliminary analysis of the histograms of the number of emergencies per day suggest that these variables follow a Poisson process (Appendix D). Therefore, non-parametric tests are appropriate. For this reason, a Kruskal-Wallis test is conducted. The results are presented in Table 7. Since all p-values are smaller than the standard 0.05 threshold (corresponding to a 5% significance level), it is possible to conclude that there is a statistically significant difference between the demand in different months and shifts.

Table 7 - Results of Kruskal-Wallis tests to the average daily volume of calls grouped by month and shift.

City	Priority	p-value	City	Priority	p-value
Lisbon	**P1**	< 2.2e-16	**Setúbal**	**P1**	0.008062
	P3	4.907e-10		**P3**	8.634e-06

6.5.2 MODEL FITTING AND FORECASTING

As mentioned in Chapter 4, methods to model EMS demand include probability distributions, time-series models, spectral analysis or neural networks. Using neural networks is not possible due to data scarcity, which would be insufficient to develop the training sets required by machine learning algorithms. Similarly, time-series models require longer series and demographic and economic factors for regressive models are not available at the required aggregation level.

Besides, by analysing the elapsed time between consecutive calls, it is possible to conclude that they closely resemble an Exponential distribution, as exemplified in Figure 20.

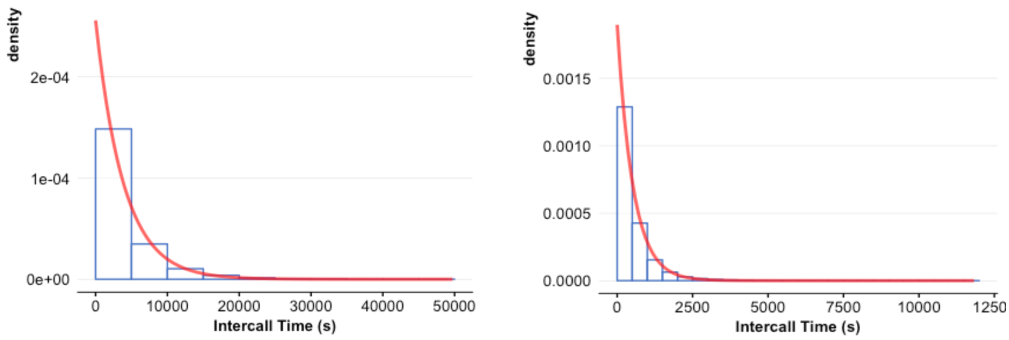


Figure 20 - Histogram and fitted exponential distribution to the elapsed time between consecutive emergency requests (Right: P1 emergencies, Lisbon; Left: P3 emergencies, Lisbon).

To verify this hypothesis, a Chi-Squared Goodness-of-Fit test is conducted to the inter-call times of P1 and P3 calls in Lisbon and Setúbal. The results are presented in Table 8.

Table 8 - Results of a Chi-Squared Goodness-of-Fit test to the fit of an exponential distribution to inter-call times.

City	Priority	p-value	City	Priority	p-value
Lisbon	**P1**	0.5616801	**Setúbal**	**P1**	0.4028691
	P3	0.6892617		**P3**	0.2540842

Given that all p-values are above the 0.05 significance level, the hypothesis that the inter-call times are consistent with an exponential distribution cannot be statistically rejected. In light of these results, emergency requests are modelled as a non-stationary Poisson process, since the arrival rate fluctuates during the day. To simplify the analysis, the Poisson process is assumed to be decomposable in three shifts and twelve months. In each of these periods, calls occur at a fixed rate, determined by fitting an

Exponential distribution to the inter-call times. As such, there are 36 different arrival rates for each call priority and region. The corresponding rates (calls/hour) are presented in Table 30 of Appendix E.

Subsequently, the fraction of emergencies coming from each region is calculated. Finally, the number of emergencies calls on each shift is calculated by using a Poisson distribution with the fitted rate parameter. Also, upper and lower bound estimates by calculating the 95% confidence intervals can be computed. As an example, the final estimates of daily call volumes for two clusters of the 1% cluster partition of Lisbon are presented in Table 9.

Table 9 - Example of the estimates of expected daily call volumes for the 1% cluster partition of Lisbon.

			Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
CLUSTER 9	**P1**	**Morning**	1.07	1.10	0.99	0.64	0.83	0.97	1.02	1.11	1.36	0.67	0.83	1.45
		Evening	0.81	0.93	0.62	0.61	0.50	0.71	0.60	0.36	0.67	0.85	0.68	0.82
		Night	0.38	0.45	0.33	0.47	0.49	0.25	0.64	0.24	0.55	0.36	0.07	0.27
	P3	**Morning**	7.78	6.74	8.28	7.65	7.69	8.28	7.76	6.86	6.00	7.87	7.46	6.42
		Evening	5.72	6.08	6.35	4.88	6.04	4.92	4.69	4.19	5.52	5.62	6.42	6.26
		Night	2.59	2.68	2.84	2.37	2.53	2.28	1.89	3.80	3.28	2.43	2.51	2.64
CLUSTER 17	**P1**	**Morning**	0.77	1.56	1.44	0.67	0.96	1.02	1.09	1.03	1.23	1.08	1.35	1.56
		Evening	0.89	0.90	0.88	0.90	1.15	1.04	0.95	0.43	0.75	0.86	0.57	1.74
		Night	0.46	0.82	0.49	0.47	0.27	0.69	0.25	0.52	0.61	0.26	0.36	0.28
	P3	**Morning**	9.64	9.71	8.98	8.76	9.03	9.20	9.31	7.43	8.67	9.39	10.0	9.53
		Evening	7.82	7.81	7.37	6.11	6.64	7.12	7.19	5.50	5.52	7.38	7.80	7.64
		Night	3.32	3.07	3.03	3.20	2.80	3.47	2.85	3.63	2.86	3.45	3.71	3.18

6.6 SERVICE AND TRAVEL PARAMETERS

As mentioned, historical records do not include travel time information. However, these records reflect the total service time, defined as the sum of travel and service time. In order to estimate the travel time, the empirical model of Budge, Ingolfsson and Zerom (2010) is used. In this model, the probability distribution of travel times for emergency vehicles travelling under "lights and sirens" (as Portuguese emergency vehicles) as a function of the travel distance is described by:

$$T = m(d)e^{c(d)\times\varepsilon}$$

Where d is the travel distance, $m(d)$ is the median travel time for distance d, $c(d)$ is the centile-based coefficient of variation and ε follows a centred t-student distribution with $\tau = 4$ degrees of freedom. In order to estimate $m(d)$, the KWH function developed by Kolesar, Walker and Hausner (1975) is used:

$$m(d) = \begin{cases} 2\sqrt{d/a} & d \leq 2d_c \\ v_c/a + d/v_c & d > 2d_c \end{cases}$$

This function assumes that the vehicle accelerates at a rate a until it reaches cruising speed v_c after a distance of d_c. Furthermore, $c(d)$ is estimated as:

$$c(d) = \frac{\sqrt{b_0(b_2+1) + b_1(b_2+1)m(d) + b_2 \times m(d)^2}}{m(d)}$$

Where b_0 represents the variability at the start and end of the trip, b_1 represents short-term speed variation within a trip and b_2 measures variability from external factors. The authors use data from the city of Calgary, Alberta, to find maximum likelihood estimators for the above parameters.

In the absence of data from the case study, this empirical model is used with the original parameters. However, some modifications are implemented. The cruising speed, v_c, is estimated as $50\ Km/h$ during the morning, $60\ Km/h$ during the evening and $70\ Km/h$ for the night. This value is different from the $v_c = 100.7\ km/h$ used by the authors in their case study, which is unreasonable for Lisbon and Setúbal where the speed limit is 50 Km/h. Nevertheless, it has been recognized that emergency vehicles travel at a faster speed, especially in less congested shifts. Furthermore, the average acceleration and deceleration, a, is estimated at $20\ Km/h/min$, about half of the $a = 41.0\ km/h/min$ of the original study.

In order to obtain estimates of travel time, the road distance between each station and postal code is determined using Google Distance Matrix API. This procedure implies that emergency vehicles are assumed follow the shortest path, since this is the distance returned by the API. The travel mode is set to "driving" and no restrictions are imposed on the route. Subsequently, the empirical model can be applied by running a Monte Carlo simulation with 4000 replications (which enables an estimate of $\pm 0.1\ min.$ with 95% confidence), implemented in a R. Depending on the trip's shift, different parameters are used to capture fluctuating traffic dynamics during the day. As an illustration, Figure 21 shows the estimated mean travel times as a function of distance for different cruising speeds, v_c. The maximum travel time, which is also a model parameter, is determined for different cluster partitions, as presented in Table 31 of Appendix E.

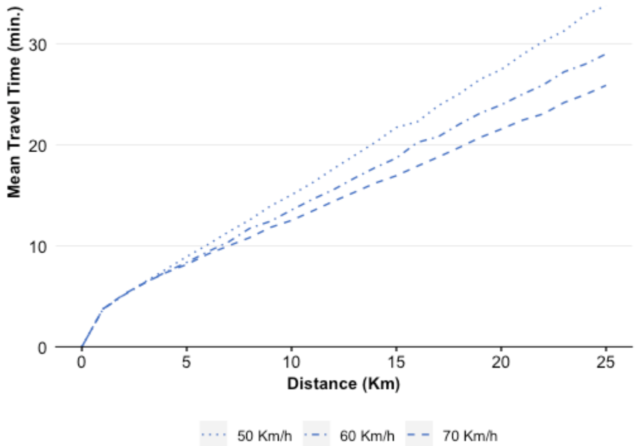


Figure 21 - Average travel time for different cruising speeds as a function of travel distance.

It is important to highlight that, given that all considered vehicles are 4-wheel vehicles, it is assumed that they have similar travel times and coverage probabilities. If MEMs had been considered, then their shorter travel time (and higher coverage probabilities) could have been accounted for by increasing their cruising speed and average acceleration in the travel time model of section 6.7.

Finally, service time is calculated by subtracting the estimated travel time from the total service time and calculating the sample mean for each demand zone, priority, care level, month and shift, using R's package *dplyr* (Wickham *et al.*, 2019). For some demand areas, no emergencies are recorded. In these cases, the service time is estimated as the global average service time for that priority, care level, month and shift. A sample of estimated service times for Setúbal is presented in Table 32 of Appendix E.

6.7 LEGISLATION RESTRICTIONS

As mentioned in Chapter 3, Portuguese legislation sets rules regarding the minimum number of vehicles that must be positioned in hospitals, depending on their emergency department category. In particular, all SUP (Multipurpose Urgency Services) and SUMC (Medical-Surgical Urgency Services) hospitals should be assigned a VMER, while SUB (Basic Urgency Services) should have a SIV.

Data from the Portuguese NHS shows that all hospitals with emergency departments in Lisbon are SUP, while Setúbal has one hospital with SUMC. Therefore, a VMER should be assigned to all these hospitals. However, one hospital in Lisbon - *Maternidade. Alfredo da Costa* - has been excluded by INEM due to lack of space for an emergency vehicle. Additionally, two other hospitals in Lisbon are not equipped with a VMER, as the legislation requires, due to insufficient vehicles. Consequently, the minimum number of vehicles and the current situation can be summarised as in Table 10.

The legislation also defines which vehicles can be positioned in each type of station. These requirements are already reflected on the allowed allocations of Table 6. Additionally, the legislation also states that one PEM must be located in each municipality and that AEMs must be located in regions where SUP or SUMC hospitals exist. However, this is not an issue in any of the study areas since these requirements are already met with the available vehicles.

Table 10 - Minimum and current number of vehicles in emergency stations.

		Required Vehicles	Current Vehicles	
	Stations	**VMER**	**VMER**	**AEM**
LISBON	Hospital São José - Lisboa	1	1	0
	Hospital Santa Maria	1	1	0
	Hospital São Francisco Xavier	1	1	0
	Maternidade Dr. Alfredo da Costa	1	*No conditions*	
	Hospital D. Estefânia	1	0	0
	Hospital Egas Moniz	1	0	1
SETÚBAL	Hospital São Bernardo	1	1	0

6.8 OBJECTIVE FUNCTION DATA

6.8.1 COVERAGE PROBABILITIES

In order to estimate the coverage probabilities, in the absence of travel time data, the empirical model described in section 6.7 is employed. As mentioned, INEM sets a coverage target of 15 minutes for all call priorities and all care levels in an urban setting. Therefore, the probability that a vehicle can cover a given emergency corresponds to the probability that it can reach the scene in 15 minutes or less. This probability is estimated by the same procedure described in section 6.7. Firstly, the distance between stations and the centroid of each area is determined using the Google Distance Matrix API. A Monte Carlo simulation is used to get a numerical estimate of the probability $P(T \leq 15\, min.)$.

Figure 22 shows the coverage probability as a function of distance for different cruising speeds. In practice, the empirical model is used to develop a gradual coverage function which depends on distance and speed (which, in turns, depends on the time of the day). As an example, Table 33 of Appendix E shows the coverage probabilities for Setúbal for 3 regions of the 1% partition and 3 stations during the morning. As in section 6.7, it is assumed that the coverage probability is independent of vehicle type.

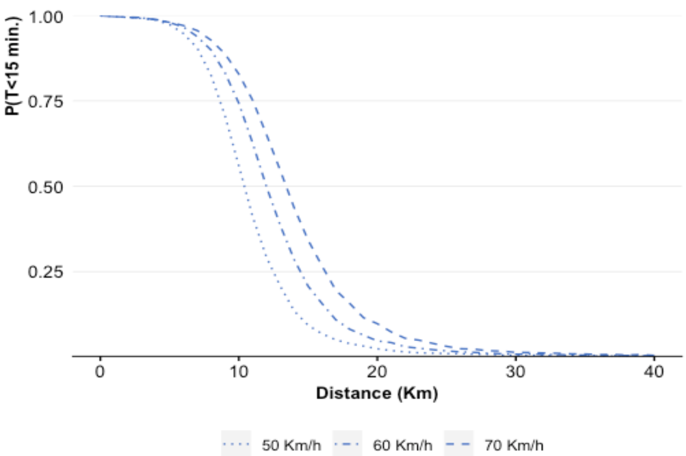


Figure 22 - Probability of travel time bellow 15 minutes for different cruising speeds as a function of distance.

6.8.2 COVERAGE WEIGHTS

Objective function weights are determined in collaboration with INEM using a Swing Weighting Procedure (Von Winterfeldt and Edwards, 1993). This method is chosen for its simplicity. Since these weights are related with covering different emergencies in different periods and shifts, a lower reference level is defined as a solution that does not cover emergencies of a given priority with a given care level, while an upper reference level is defined as a solution providing full coverage to those calls.

To assess the W1 weights, the decision-maker is asked to consider the improvements from the lower reference level (no coverage) to the upper reference level (full coverage) on each priority/care-level combination. The most important swings are set as providing ALS and BLS to P1 calls, which the decision-maker considers to be equally attractive. The remaining swing is ranked as the less preferred swing. Subsequently, the decision-maker is asked to compare the other swings to the most preferred swings, for which a value of 100 is established, by comparing the improvement in the attractiveness of the solution. After quantifying the swings, the weights are obtained via normalization (Table 11).

Table 11 - Objective function weights (W1).

	P1		P3
	ALS	**BLS**	**BLS**
W_1	1	1	0.75

Repeating this procedure for W2 and W3 weights, lower and upper reference levels are set as no coverage and full coverage of all emergencies on each month and shift, respectively. The decision-maker considers all to be equally attractive. As such, all W2 and W3 weights are set to 1.

6.8.3 SYSTEM COSTS

As mentioned, the GPCG is not responsible for cost analysis at INEM. Therefore, cost information cannot be made available. Nonetheless, given the importance of considering costs, these are estimated based on public information. In particular, the 2016 Accounting Report (Instituto Nacional de Emergência Médica, 2016a) is used, since it is more detailed that subsequent versions. In order to simplify the analysis, cost categories are developed. These categories are developed by adapting the cost framework proposed by Lerner *et al.* (2007) to INEM's reality. Only costs which are directly or indirectly affected by the vehicle location decisions are considered. A summary of considered cost categories together with the model parameter to which they belong is presented in Table 12.

Table 12 - Considered cost categories from Lerner et al. (2007) and corresponding model inputs.

Cost category		Description	Parameter
Human Resources	Salaries	Compensation paid to TEPHs.	$OperatingCost$
	Overtime	Shift bonus.	$OperatingCost$
Physical Station	Acquisition	Station installation (e.g. administrative costs).	$OpeningCost$
	Operation	Rent for AEM. PEM and VMER subsidies.	$CapacityCost$
	Replacement	Station decommissioning.	$ClosingCost$
Vehicles	Operation	Fuel (AEM, VMER and SIV) and exit prizes (PEM, RES and NINEM).	$AssignmentCost$
	Maintenance	Repairs and others for AEM, VMER, SIV. Included in exit prizes/subsidies for PEM, RES, NINEM.	$OperatingCost$
Consumables	Acquisition	Only for AEM. Other vehicles are resupplied by partner.	$AssignmentCost$

Although the model allows for costs to depend on both period and shift, it is assumed that all cost categories are independent of time and shift considered. While base salaries and exit prizes are not expected to fluctuate during the planning horizon, additional benefits may depend on the working shifts. Additionally, cost for RES are excluded because neither of the two regions uses these vehicles.

Opening and Closing Costs

Since INEM uses existing public facilities as stations, there are no construction or decommissioning costs. Nevertheless, there are administrative costs related with opening and closing facilities. As public reports do not include information to estimate these costs, they are assumed to be negligible.

Capacity Costs

Capacity costs are monthly or quarterly fixed costs paid for each vehicle housed at a station. For VMER, the legislation establishes an allowance of $6.800\ €/month$ to the base hospital. For PEM, a quarterly allowance is provided depending on the total number of dispatches per month. In both Lisbon and Setúbal, since PEMs are not dispatched on average more than 100 times per month (which is the limit for the lowest category of subsidies), the allowance is $6.400\ €/quarter$. Additionally, for AEM and SIV, capacity costs are related with the rent required by the receiving public institution. Unfortunately, these rent values are not available, so these are assumed to be negligible. For NINEM, there are no capacity costs, as these are covered by higher exit prizes.

Table 13 - Capacity cost per vehicle per day (€).

Capacity Cost	AEM	SIV	VMER	PEM	NINEM
Health Units	0	0	226.67	—	—
INEM HQ	0	0	—	—	—
Firefighters/RC	0	—	—	71.11	0
Police/Schools	0	—	—	—	—

Operating Costs

Operating costs include crew salaries and vehicle repairs. It is assumed that salaries are independent of shift. Salaries are paid only for AEM and SIV crews, which are staffed by INEM employees. NINEM, RES, PEM and VMER crews are paid by other entities. A TEPH's salary is $738.05\ €/month$ while an INEM nurse receives $1201.48\ €/month$, with an additional 25% bonus due to shift work, and work 35 hours per week (República Portuguesa, 2016, 2019). Two TEPHs operate an AEM, and one TEPH and one nurse operate a SIV. Therefore, computing hourly salaries, the cost of operating each vehicle can be computed. Additionally, repair cost can be estimated by taking into account that, in 2016, INEM spent $4.442.616\ €$ in fleet repair and maintenance. Dividing this amount by the total number of vehicles, 611 (excluding outsourced Helicopters and NINEMs), the days in the year and the number of shifts, the average repair cost per vehicle per shift can be estimated. It is assumed that all vehicles have similar maintenance costs and that these are distributed evenly throughout the year and shifts. The final cost estimates per vehicle per shift are presented in Table 14. NINEMs are omitted since their cost is 0.

Table 14 - Estimated cost for INEM per vehicle per shift (€).

Vehicle	Salaries	Shift	Repairs	Total
PEM	0.00	0.00	6.64	**6.64**
AEM	74.98	18.74	6.64	**100.36**
VMER	0.00	0.00	6.64	**6.64**
SIV	98.52	24.63	6.64	**129.78**

Assignment Costs

Finally, assignment costs depend on the emergencies assigned to a vehicle. For INEM vehicles, this cost comprises fuel and medical consumables. For RES, PEM and NINEM, INEM pays an exit prize, which depends also on the location of the emergency. VMER's assignment costs are supported through the monthly subsidy. For simplicity, average fuel costs per dispatch are assumed to be equal for all vehicles. These can be estimated by dividing total fuel costs in 2016, $1.167.927\,€$, by the number of INEM vehicles dispatches, 218.743. A similar procedure is applied to determine average costs of medical consumables per dispatch, which amount to a total cost of $570.865\,€$. Furthermore, Table 15 presents the three first distance classes and corresponding exit prizes for PEM/RES (Associação Bombeiros para Sempre, 2015). It is assumed that both ambulances are always manned by at least one TAS (*Tripulante de Ambulância de Socorro*)[3], as required by the legislation.

Table 15 - Exit prizes paid to PEM and RES (€/dispatch) (Associação Bombeiros para Sempre, 2015).

Distance	PEM	RES
0 – 15 Km	5.8	11.33
16 – 40 Km	11.0	18.54
41 – 65 Km	21.6	31.93

Using this information and the estimated distances between stations and each demand areas, assignment costs of PEMs can be calculated. Regarding NINEMs, since there are no formal agreements, exit prizes are not regulated nor publicly available. Therefore, it is assumed that these costs are similar to RES ambulances. The estimated assignment costs are presented in Table 16.

Table 16 - Estimated assignment costs (€/dispatch).

Vehicle	Fuel	Consumables	Exit Prize	Total
AEM	5.34	2.61	0	7.949
SIV	5.34	2.61	0	7.949
PEM	0.00	0.00	5.8/11	5.8/11
RES	0.00	0.00	11.33/18.54	11.33/18.54

6.9 OTHER PARAMETERS

Additional model parameters are also estimated from interviews:

- A station/vehicle pair is responsible for a region if it covers at least 15% of its calls ($\varepsilon = 15\%$);
- There must be 2 vehicles responsible for providing each care-level to each call ($N = 2$).
- A reliability level for vehicle availability is set at 80%. Therefore, $\rho_{max} \leq 0.37015$.
- There are no restrictions on the assignment of stations to calls, thus $\delta_{fd} = 0, \ \forall f \in F, d \in D$.
- All stations, if opened and selectable, must remain in operation for at least one month;
- Since the model is applied to a year, the effects of inflation are neglected ($i^t = 0, \ \forall t \in T$.
- All vehicles capable of providing BLS can provide it to P1 or P3 calls, which defines parameter γ_{vpl} (recall that $\gamma_{vpl} = 1$ if a vehicle of type v can provide care level l to a call of priority p).

6.10 CHAPTER CONCLUSIONS

The present chapter describes the data collection and treatment procedures required to estimate model parameters from real data. Different sources of information include interviews with INEM practitioners, historical records and public data. Applied techniques include GIS, point-aggregation algorithms, Google's Geocoding and Distance Matrix APIs and statistical analysis in R. It Is possible to conclude that accurately collecting and treating the real data is challenging given the large number of parameters required, which may not be readily available. Nevertheless, this step paramount for a successful application of the model. The next chapter describes the results of model application to the case-study.

[3] This assumption is required because PEM and RES ambulances without a TAS receive a lower exit prize. Nonetheless, in theory, all PEM and RES should have at least one TAS in their crew.

7. CASE STUDY RESULTS

This chapter presents the application of the proposed model to INEM's case study. Besides presenting computational results, the chapter focuses on discussing recommendations for INEM. It is divided in eight sections. Section 7.1 describes model implementation and validation. Section 7.2 focuses on a preliminary scalability analysis of the model, while section 7.3 introduces the computational experiments to be conducted. Section 7.4 presents an analysis of the current system and sections 7.5 and 7.6 describe the results of the experiments, leading to a set of recommendations for the CPCG in Section 7.7. Section 7.8 presents a sensitivity analysis to several model parameters. Finally, section 7.9 presents the computational results of the proposed heuristic. Conclusions are postponed to Chapter 8.

7.1 MODEL IMPLEMENTATION

In order to apply the model to INEM's vehicle location problem, the model is implemented in IBM ILOG CPLEX Optimization Studio 12.8.0.0, using the IDE and Optimization Programming Language (OPL). Data from the case study is initialized externally, being inputted through Excel. Slicing is used to reduce the computational burden by determining the valid combinations of parameters *a priori* (IBM Corporation, 2010). Furthermore, Visual Basic for Applications (VBA) in Excel is used to generate the allowed combinations of indices. The heuristic approach is implemented in OPL script, a Java extension developed specifically to control CPLEX models, which is available in the CPLEX Optimization Studio IDE (IBM Knowledge Center, 2019a).

A Lexicographic Ordering approach is used to handle multiple objectives. Objectives are ranked according to their priority for INEM in decreasing order of importance: Z1, Z2 and Z3. Three runs are conducted for each experiment. Firstly, coverage is maximized (run 0). Subsequently, this objective is bounded by a constraint (lower bound) and costs are minimized (run 1). In the third run, an upper bound on costs and a lower bound on coverage are set and equity is maximized by minimizing $Z3$ (run 2). The solution of the previous run is used as a MIP start in CPLEX, given that it is always a feasible solution.

All experiments are conducted on a 2.40 GHz Intel Core i7-4700MQ processor and 12.0 GB of RAM laptop running Windows 10. Hereafter, the exact parameters presented in the previous chapter are referred to as "base line scenario". In these scenarios, level 1 potential stations are considered, since they are preferential for INEM. Level 2 stations are included only when stated.

7.2 MODEL VALIDATION AND SCALABILITY ANALYSIS

7.2.1 MODEL VALIDATION

Validating the proposed model before carrying out further analyses is paramount to ensure its accuracy in representing the real system, guaranteeing that any simplifying assumptions and corresponding mathematical formulations developed during the modelling stages adhere to reality (Williams, 2013).

In order to validate the proposed model, two small instances are firstly analysed. The first instance includes three emergency stations, demand nodes, shifts and vehicles, and two planning periods. The second instance consists of five emergency stations, demand nodes and time periods, and four vehicles. On both instances, extreme conditions are introduced by varying several parameters and testing the model's response to such changes, in order to ensure that the formulated constraints generate the expected behaviour. For instance, by setting the number of available vehicles to zero, total coverage drops to zero. On the other hand, if facility relocations are not allowed (the maximum number of closed/opened stations is set to zero), the station configuration remains unchanged, but vehicles are reallocated to improve coverage. Furthermore, by varying coverage probabilities, the configuration of stations is adjusted, favouring stations with higher coverage probabilities, and by allowing more station relocations, total coverage increases.

7.2.2 SCALABILITY ANALYSIS

Before proceeding to apply the model, it is important to assess how it behaves when the instance size increases, so that the experiments can be conducted in a cost-effective manner. In order to assess the scalability of the model, several instances of different sizes (corresponding to different cluster partitions) are analysed. Details on instance size and scalability analysis results are presented in Appendix F.

From the results, it can be concluded that the computational effort increases exponentially with the instance size. This is mainly due to the fact that, being a dynamic model, the number of variables and constraints increases significantly even for small networks. In fact, by considering 12 time periods and three shifts, the number of variables (indexed by these parameters) increases 36 times when compared to a static model. The results also show that proving optimality is challenging even though the solution becomes near-optimal quickly. Furthermore, the best bound is close to the optimal solution, suggesting that the linear relaxation is relatively tight. Furthermore, the computational burden increases considerably in the second and, mostly, third run of the Lexicographic method. This is because additional constraints related with the previous objectives render the model much more difficult, Nevertheless, given that speed is not crucial in this study, CPLEX is used except when stated otherwise.

The results also suggest that the chosen cluster partition only slightly influences the results. In the tested instances, using larger aggregation areas results in an overestimation of coverage of 2.38% in Lisbon and 5.86% in Setúbal. The equity objective presents an underestimation of 4.78% in Lisbon and 2.07% in Setúbal. Regarding costs, the results are not conclusive. In Lisbon, they are underestimated by 1.40% in Lisbon, while in Setúbal they are overestimated by 8.37%.

In light of these results, and in order to reduce the computational cost and allow for more experiments to be carried out, the remaining experiences are conducted using the smallest instances for Lisbon. For Setúbal, the largest instances are used since the computational cost is still acceptable. A relative gap of 0.5% is used when maximizing coverage, while a gap of 5% is applied for the remaining objectives. Additionally, a time limit of 24 hours is set for each run. The following sections focus on exploring the model to assess the impact of alternative policies in the SIEM's expected performance. Computational times for all the experiments are detailed in Appendix H.

7.3 CURRENT SOLUTION PERFORMANCE

Before proceeding to study alternative system configurations, it is important to assess the current system performance. For this purpose, a heat map of P1 and P3 emergencies in Lisbon and Setúbal from 2017 and the first semester of 2018 is presented in Figure 23, together with the positions of existing stations classified as having ALS or BLS vehicles.

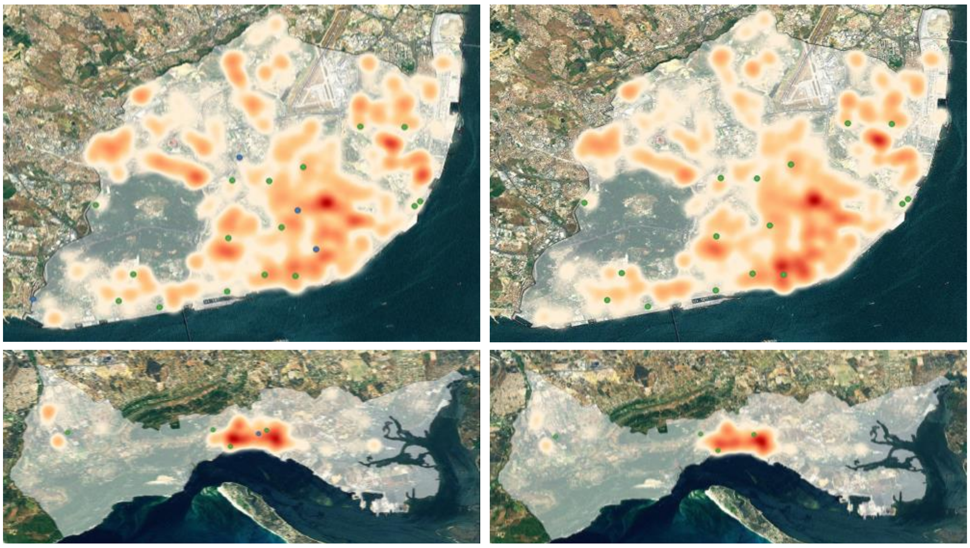


Figure 23 - Heatmaps of the current system and current emergency stations (Top: Lisbon; Bottom: Setúbal; Left: P1 calls, BLS and ALS stations; Right: P3 calls and BLS stations).

A qualitative analysis of Figure 23 suggests a higher emergency density in the centres of both cities, where population density is higher due to residents, commuters and tourists. As would be expected, more stations are located in these areas. Nevertheless, the north of Lisbon and a large part of Setúbal are apparently less covered. It is also useful to analyse heatmaps of the activity areas of each emergency station and the distribution of emergency requests for the morning shift (Table 17 and Table 18), the only where the entire fleet operates and, as such, can be used for comparison.

Table 17 - Distribution of emergency requests per vehicle during the morning shift for Setúbal.

Vehicle	P1	P3	Aggregate	Vehicle	P1	P3	Aggregate
BV Setúbal	16.24%	27.87%	25.71%	**Outside Vehicle**	10.53%	15.96%	14.96%
AEM Setúbal 1	14.25%	24.44%	22.55%	**VMER Setúbal**	40.69%	0.04%	7.59%
AEM Setúbal 2	14.18%	23.57%	21.83%	**CVP Setúbal**	4.11%	8.11%	7.37%

Table 18 - Distribution of emergency requests per vehicle during the morning shift for Lisbon.

Vehicle	P1	P3	Aggregate	Vehicle	P1	P3	Aggregate
BV Beato e Olivais	4.06%	12.00%	10.61%	**BV Ajuda**	1.37%	4.11%	3.63%
Outside Vehicle	6.24%	10.55%	9.79%	**AEM Lisboa 13**	2.38%	3.88%	3.62%
BV Lisbonenses	2.52%	6.32%	5.66%	**AEM Lisboa 5**	2.41%	3.71%	3.48%
BV Lisboa	1.60%	5.84%	5.10%	**VMER S. José**	18.68%	0.03%	3.28%
BV Cabo Ruivo	1.80%	5.40%	4.77%	**AEM Lisboa 7**	1.81%	3.53%	3.23%
AEM Lisboa 1	2.85%	4.72%	4.40%	**AEM Lisboa 6**	1.97%	3.49%	3.23%
AEM Lisboa 10	2.89%	4.69%	4.38%	**AEM Lisboa 12**	1.61%	3.49%	3.16%
AEM Lisboa 2	2.54%	4.57%	4.22%	**SIV Lisboa**	15.03%	0.06%	2.67%
AEM Lisboa 9	2.53%	4.50%	4.16%	**VMER Sta. Maria**	13.02%	0.01%	2.28%
BV Campo de Ourique	1.64%	4.64%	4.12%	**RSB Lisboa**	0.48%	1.09%	0.98%
AEM Lisboa 11	2.35%	4.43%	4.07%	**VMER SF Xavier**	5.16%	0.01%	0.91%
AEM Lisboa 4	2.46%	3.99%	3.72%	**AEM Lisboa 14**	0.37%	0.62%	0.58%
AEM Lisboa 3	2.10%	4.06%	3.72%	**AEM Lisboa 15**	0.13%	0.25%	0.23%

Regarding Setúbal, most emergencies are concentrated on the city centre and, as such, vehicle activity patterns are more balanced. Workload is approximately distributed, except for CVP Setúbal, that receives significantly less calls, which could indicate that these vehicles could be located elsewhere. External dependency is high (close to 15%). Regarding Lisbon, external vehicles operate mainly on the northern part of the city. Additionally, vehicles with higher dispatching fractions are located closer to the city centre and the eastern part of the city, suggesting that providing additional vehicles to these regions (or relocating vehicles from other regions) may be beneficial. NINEMs present the highest dispatch fractions, which is surprising given that INEM favours their own vehicles and those with which protocols exist (PEMs). Some AEMs present low dispatch fractions, which suggests that their positions can be improved. Furthermore, the external dependency is also significant, reaching close to 10%.

To assess the performance of the current system under the proposed model, the current solution is inputted (through the sh variables) and the value of different objectives are retrieved. This requires assuming that the allocation of requests to vehicles is optimal. The solutions are presented in Table 19. These values are used for comparison with alternative policies.

Table 19 - Estimated performance of the current system using the optimization model.

Instance	**Coverage** (Z1)	**Gap (%)**	**Cost** (Z2)	**Gap** (%)	**Equity** (Z3)	**Gap** (%)
L0.1	1686.135	0.10%	66494.661	0.01%	17.856	4.86%
S2.1	320.430	0.00%	17119.534	0.00%	18.321	0.48%

Note that, under the demand parameters of the previous chapter, the theoretical maximum expected coverage (accounting for the objective function weights) is 2032.20 in Lisbon and 333.79 for Setúbal. Therefore, the current system can cover around 83% and 96% of the calls, respectively.

7.4 BASE SCENARIO OPTIMIZATION

In this section, improvements for the current system under the full set of restrictions presented in previous chapters are studied. For this purpose, the base-line optimization instances are used again and solved in CPLEX. The resulting objective function values are presented in Figure 24, as well as the corresponding variations with respect to the current system. Recall that a negative variation in the equity indicator is desirable, as it means a reduction in the maximum average travel time. In this experiment, only seasonal stations are considered. The effect of seasonal PEMs is analysed separately, given INEMs interest in studying the impact of this policy.

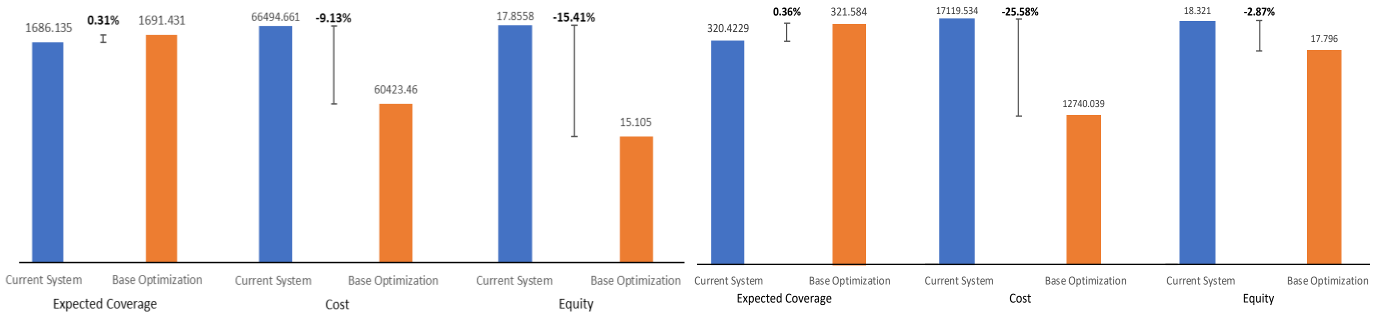


Figure 24 - Comparison of objective function values of the base-line optimization scenarios against the current system performance (Left: Lisbon; Right: Setúbal).

The results highlight that, under the current restrictions imposed by INEM and the legislation, the present vehicle configuration can only be slightly improved in what concerns the primary objective, expected coverage. In fact, by analysing the resulting solutions (and comparing them with the existing configuration), it is possible to verify that only limited changes are possible.

This is mainly because all stations are classified as non-selectable, meaning that they cannot be relocated. Since the maximum number of stations in most months is also fixed to the current number of stations (20 in Lisbon and 5 in Setúbal), the underlying structure of the emergency system network cannot be modified. As a result, only seasonal stations can be opened during the summer months.

In Lisbon, seasonal stations include *Centro de Saúde de Alvalade* (June), *Centro de Saúde de Benfica* (July), *Centro de Saúde da Graça* (July, August), *Centro de Saúde de Marvila* (August) and *Hospital de Cruz Vermelha* (September). In Setúbal, one seasonal station is allowed,: *Centro de Saúde de São Sebastião* (June) and *Centro de Saúde do Bonfim* (July).

Figure 25 displays the proposed seasonal stations in Lisbon, showing that these sites are mainly located in areas where there is a considerable density of emergencies without stations nearby or, alternatively, near high emergency density areas. These results agree with what would be expected from the analysis of the previous section.

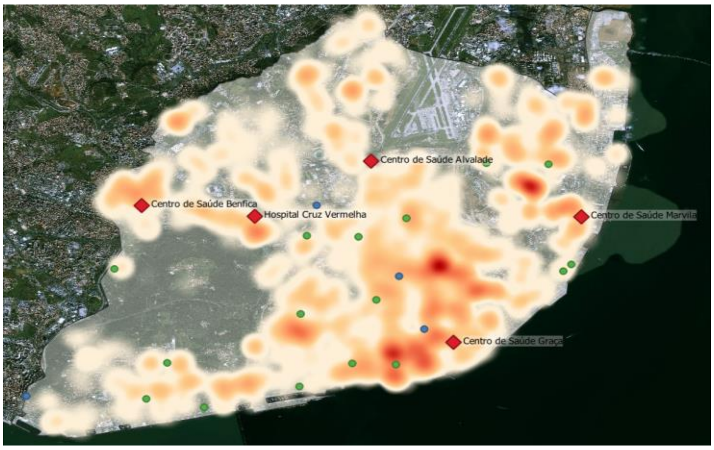


Figure 25 - Proposed seasonal stations in Lisbon.

Additionally, emergency vehicles are relocated to improve the system's performance. Yet, some vehicle types (NINEM and PEMs) cannot be relocated as they are non-selectable. This is particularly limiting in Setúbal, where only three vehicles are selectable (VMER and two AEMs), and, as such, are the only whose position can be adjusted. As a consequence, the only proposed modification is the relocation of one AEM to the seasonal station during the summer and to *Hospital de Setúbal* during the remaining months. The other AEM is forced to remain at its original station, which is non-selectable and, therefore, must remain open with, at least, one vehicle. The VMER cannot be moved due to legislation.

On the other hand, in Lisbon, a significant proportion of vehicles are relocated, since a large part of the fleet is composed of selectable vehicles. In particular, AEMs are relocated to improve the system's performance. For instance, by May, two AEMs are moved to *INEM Sede*, while the SIV is moved to *INEM – Rua Infante D. Pedro*, which also keeps one AEM. Conversely, two AEMs from *Centro de Saúde Olivais* are allocated to *Hospital Santa Maria, Hospital de São José* and, finally, one AEM from *Hospital Curry Cabral* is located at *BV Beato*. These relocations are in line with what would be expected given the considerations of the previous chapter. *Hospital Santa Maria* is the closest station to the Northern part of the city, which is considerably uncovered. *BV Beato*, which is the vehicle with highest fraction of dispatches, is reinforced with one AEM. *Hospital de São José* lies at the heart of the city centre, where the density of emergency calls is higher. The SIV is moved to *INEM – Rua Infante D. Pedro*, which is slightly closer to the centre of the city as well. When seasonal stations are enabled, AEMs are repositioned to these stations. VMERs are maintained at *Hospital de Santa Maria, Hospital de São José* and *Hospital São Francisco de Xavier*, because, again due to legislation, these stations must keep at least one VMER and there are only three available.

It is interesting to note that although the improvement in expected coverage is modest, the reduction in cost is more significative: around 9.13% for Lisbon and 25.58% for Setúbal. The same result is valid for the equity objective, where the results show a reduction in the maximum average travel time in the system of 15.41% in Lisbon and 2.87% in Setúbal. Therefore, it is possible to conclude that, although the current restrictions are considerably tight and do not allow many modifications of the system, attractive improvements can be attained in what concerns cost and the equity, together with a modest improvement in the SIEM's ability to cover emergencies. These improvements are mainly due to the possibility of relocating some vehicles (mostly AEMs) and opening seasonal stations during the summer.

As an example, Appendix G shows the proposed solution for Lisbon. As expected, the model outputs the number of vehicles which are to be located on each station and on each shift. The main difference between the solution of Table 36 and the current solution is that, besides indicating which stations are good candidates to be seasonal stations, AEMs are progressively adjusted to better fit the demand requirements. However, these adjustments are not abrupt and, from one month to the next, only slight changes are introduced. Conversely, the model also suggests varying the working shifts (even without relocating vehicles) as a means of improving coverage. Although this would require more flexibility from INEM, it could be an effective approach to increase performance without structural changes to the system. In Setúbal, the solution is practically identical to the current system because, as explained, the imposed constraints do not enable many modifications to be carried out.

7.5 ALTERNATIVE POLICY TESTS

Given that the current restrictions prevent the attainment of a more effective solution, it becomes important to analyse alternative policies which might improve the SIEM's performance. The following sections discuss several experiments that seek to identify promising courses of action. Firstly, current practices are assessed, namely the use of summer-seasonal PEMs. Subsequently, several fleet expansion scenarios are tested in order to understand how INEM should proceed if the current practice of incremental changes is to be continued. The following section examines the effect of considering level 2 stations as potential sites for new vehicles. Finally, the impact of legislation in concerning VMERs and four alternative operating scenarios – free-vehicle relocation green-field scenario and a fully flexible EMS – are assessed in order to show how the current system could perform if the numerous restrictions are lifted. All experiments are conducted using CPLEX.

7.5.1 THE IMPACT OF SEASONAL PEMS

Currently, INEM deploys seasonal PEMs during the summer. In order to assess the impact of this policy, an additional type of vehicle is included in the model (Seasonal PEMs, SPEM). This vehicle is similar to the PEM except that it is selectable, i.e. it can be repositioned or removed from one period to the next. The impact of this vehicle is evaluated by fixing the remaining vehicles at their original stations and enabling SPEMs to be located freely during the summer months. The results of this analysis are summarized in Figure 26.

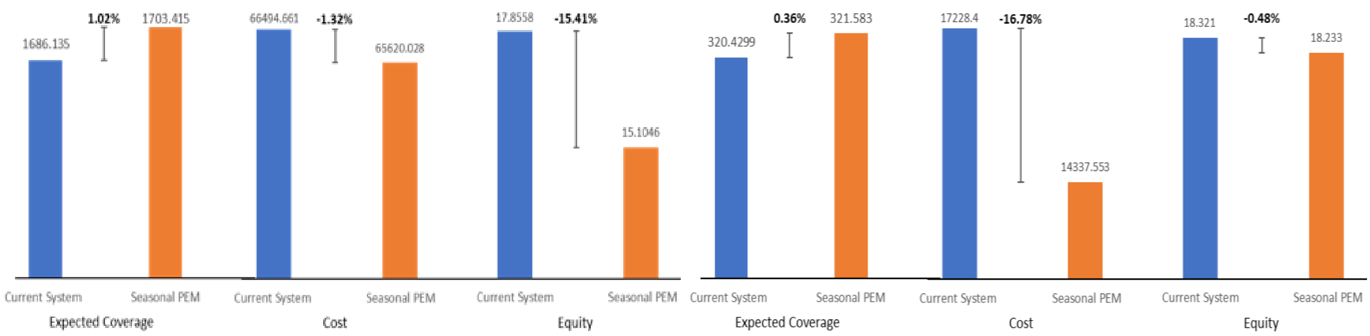


Figure 26 - Seasonal PEM results (Left: Lisbon; Right: Setúbal).

The model suggests locating Setúbal's seasonal PEM in *BV Setúbal – Azeitão*, which is, in fact, the only available location for this vehicle. In Lisbon, the seasonal PEM should be located at *BV Cabo Ruivo*. The results suggest that seasonal PEMs do contribute to improve the system's coverage by 0.20% in Setúbal and 1.02% in Lisbon and the equity indicator by 15.4.% and 0.48%, respectively.

Surprisingly, costs from the seasonal PEMs are smaller than the current system. Still, these costs are higher than the base line optimization. One reason for this is that exit prizes depend on the distance the PEM must travel. Without seasonal PEMs, some emergencies must be assigned to PEMs operating further away, since the existing PEM in this area is not capable of covering all emergencies. With an additional PEM in this region, although an associated capacity cost exists, exit prizes can be reduced due to the shorter travel distances. Additionally, by adjusting the shifts in which each vehicle operates, operating costs can also be reduced because PEMs do not have operating costs for INEM. These considerations are validated by Figure 27, which compares the different cost categories for the current system and seasonal PEM scenarios for Setúbal.

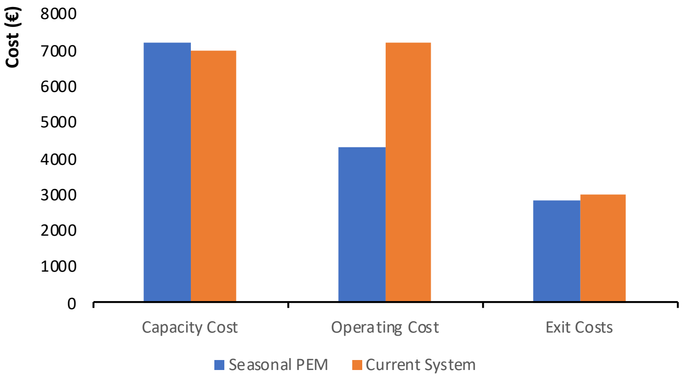


Figure 27 - Comparison of cost categories of seasonal PEMs against the current system.

It is interesting to evaluate if SPEMs are currently deployed in the most effective months or if, alternatively, they could be more beneficial in other periods. Presently, seasonal PEMs are used during three consecutive months – July, August and September. To test different alternatives, the starting month of SPEM operation is varied, and these vehicles are assumed to operate for three consecutive months. The resulting coverage values for different deployment months are presented in Figure 28.

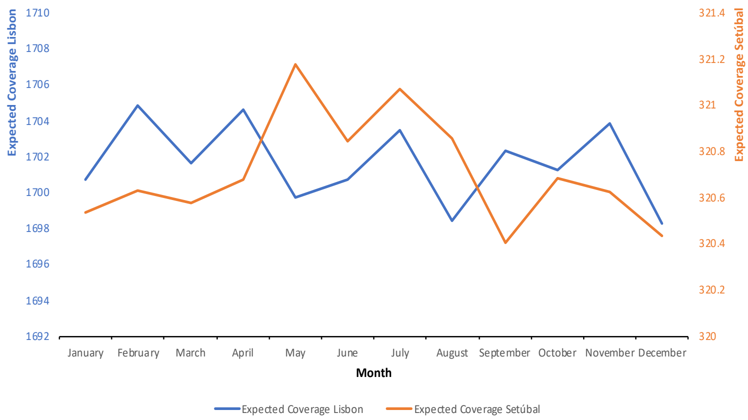


Figure 28 - Effect of different deployment months in seasonal PEM's impact.

The results indicate that, in Setúbal, SPEMs are more effective during the summer, although operating these vehicles during May, June and July may provide a slight advantage. On the other hand, in Lisbon, SPEMs would be marginally more advantageous during the first months of the year, namely February, March and April. This is in line with what would be expected from the demand analysis presented in the previous chapter, where demand is greater during the winter. One possible reason for seasonal PEMs being more advantageous during the summer is that this additional vehicle is located in *Azeitão*, near popular destinations for summer holidays.

7.5.2 FLEET EXPANSION ANALYSIS

It is also important to understand if, keeping the original restrictions, expanding the emergency fleet could result in a performance improvement of the system. This question is in line with the current planning practice at INEM, which focuses on incremental vehicle additions to the system.

For this reason, the impact of adding additional units of each type to the fleet is analysed by increasing the number of vehicles by one, two, three and four units. It is assumed that a new station is allowed for each new vehicle, in line with current practice. It is further assumed that the additional vehicle operates on all shifts. Cost is excluded from this analysis, since vehicle purchasing costs are not available, which is an important component of the expansion scenario. With these costs, it would be possible to develop a cost-benefit analysis and determine which vehicle would have a higher impact per unit cost in the system performance, thus supporting these decisions in both performance and cost metrics. The results are presented in Figure 29.

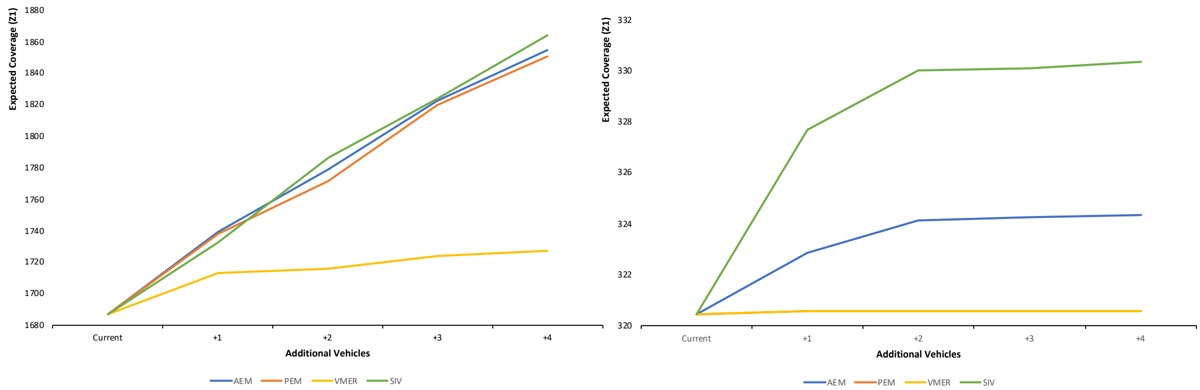


Figure 29 - Fleet expansion scenarios for level 1 stations (Left: Lisbon; Right: Setúbal).

It is possible to conclude that adding different types of vehicles results in different impacts in the system's ability to cover emergencies. In Lisbon, the results show that adding AEMs, PEMs and SIVs result in an almost equivalent performance improvement. For instance, one SIVs leads to an improvement of 2.7%, while an additional AEM leads to 3.1% increase and a PEM to 3.0% increase. Conversely, adding four SIVs results in a 10.5% coverage growth, four additional AEMs results in a 9.9% improvement and four PEMs in 9.7%. This similarity is mainly because AEMs and PEMs are capable of providing the same care levels and, given that there are many stations capable of housing both vehicles, their performance is identical. Nevertheless, AEMs still outperform PEMs for all scenarios, mostly due to their increased flexibility in what concerns the initial location and potential relocations - AEMs are selectable and can be located in many emergency stations types, while PEMs are non-selectable and limited to firefighter/Red Cross facilities. On the other hand, SIVs are differentiated vehicles, being able to provide both ALS and BLS. Despite being differentiated, this advantage is offset by the fact that the set of potential stations for SIV ambulances is more limited. Consequently, SIVs end up having similar results to AEMs and PEMs, although they have a slight edge over these alternatives. Contrarily, adding VMERs leads to a markedly lower performance improvement. This is mainly due to the fact that VMERs are less versatile and, additionally, can only be positioned at a limited set of stations.

Regarding Setúbal, SIVs are clearly the most beneficial vehicle to be added to the fleet, given their versatility. Recall that, currently, no SIV operates in Setúbal. However, the results show that providing one of these vehicles could be an interesting option, as one additional SIV leads to an improvement of 2.26%, while two, three and four additional SIVs yield improvements of 2.99%, 3.02% and 3.09%. In second place, the addition of AEMs also generates an increase in expected coverage, although relatively smaller. In particular, one AEM increases coverage by 0.76%, two by 1.16%, three by 1.20% and four by 1.22%. Lastly, additional PEMs and VMERs practically do not contribute to increase the system's ability to cover emergencies given that, currently, there are almost no emergency stations capable of receiving these vehicles. Any additional VMER would have to be located at *Hospital de Setúbal*, where one VMER is already stationed. This station can only receive one more vehicle and, as such, adding more than one VMER would not have any impact.

Therefore, if the fleet is to be expanded in one unit, one AEM should be purchased for Lisbon and one SIV for Setúbal. However, it is important to bear in mind that adding two SIVs is more beneficial in Lisbon than two AEMs. The results also suggest that, from an expected coverage perspective, expanding the fleet through the purchase of AEMs seems to be more advantageous than through PEMs. Additionally, with the current limitations in finding bases for VMERs, increasing the number of these vehicles is not significantly attractive.

Besides enabling the comparison of several expansion scenarios, the model also suggests the sites for additional vehicles. As an example, Table 20 presents the suggested stations for one additional vehicle of each type during the first month of the planning period.

Table 20 - Suggested stations for one additional vehicle of each type during the first month of operation.

Vehicle	Lisbon	Setúbal
AEM	*Hospital Pulido Valente*	*Centro de Saúde Bonfim*
SIV	*Hospital Santa Maria*	*Centro de Saúde Bonfim*
PEM	*Regimento Sapadores Bombeiros - Benfica*	*BV Setúbal - Azeitão*
VMER	*Hospital Dona Estefânia*	*Hospital de Setúbal*

The value of the equity objective (Z3) showed to be insensitive to most expansion scenarios and most fluctuations occur in steps. For instance, for Setúbal, equity is improved by 8.92% if two AEMs or VMERs are added to the fleet. If more than two of such vehicles are added, the equity objective remains unchanged. However, any other expansion scenario leads only to a modest improvement of 0.5%. Therefore, in order to improve equity, the results suggest that adding two VMERs or AEMs in Setúbal is the best choice, while for Lisbon, the addition of one VMER or SIV is the most promising option.

7.5.3 ADDITIONAL EMERGENCY STATIONS: LEVEL 2

As mentioned, level 2 stations are less preferred than level 1, since they require agreements with organisations outside the Ministry of Health. Therefore, it becomes relevant to assess if considering these sites as potential stations for new emergency vehicles is worth the additional effort. To do so, the previous experiment is replicated for AEMs, since these are the only vehicles which can be located at level 2 stations. Figure 30 compares the expected coverage for both layers of stations.

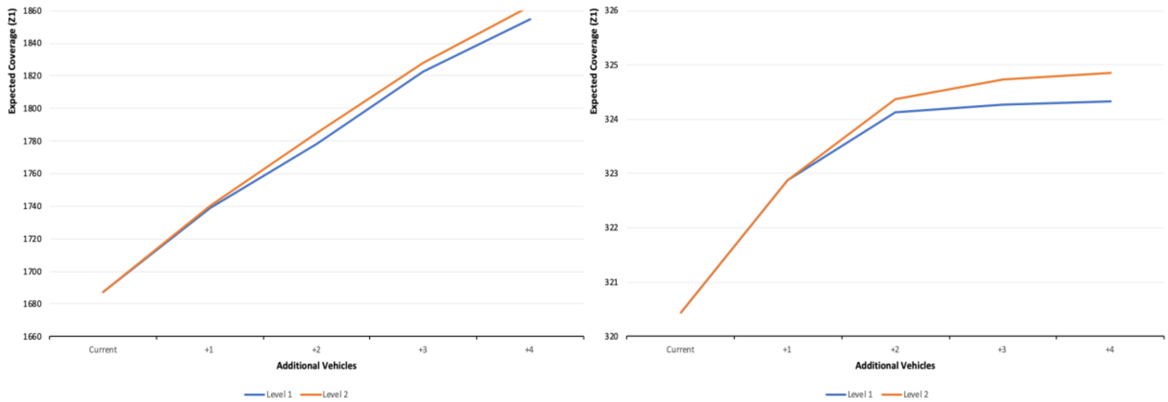


Figure 30 - Comparison of fleet expansion scenarios for level 1 and level 2 stations (Left: Lisbon; Right: Setúbal).

The results indicate that considering level 2 stations provides only marginal benefit to the system's performance when compared to level 1 stations. This is mainly due to the fact there is always a level 1 or existing station close to any level 2 station and, as such, they provide only a marginal improvement over considering a smaller subset of potential stations. The equity objective, on the other hand, is considerably reduced with level 2 stations, decreasing from around 16.99 minutes to 13.59 minutes.

Nevertheless, the proposed solutions under this scenario do take advantage of level 2 stations. For instance, one additional AEM in Setúbal, considering level 2 stations, should be located in in *PSP Setúbal – 2ª Esquadra*, while one additional AEM in Lisbon is suggested to be located at *PSP Benfica – 20ª Esquadra* and *Escola Básica Alta de Lisboa*, depending on the month.

Note that the benefit of considering level 2 stations seems to increase as the number of vehicles added to fleet also increases. This may suggest that, although in a short-sighted perspective, considering level 2 stations may not lead to great benefits, in the long run, using these stations may provide improvements in the system's performance. Therefore, if it is expected that only few vehicles will be added to the fleet in the future, considering level 2 stations may not be required. However, if more vehicles are to be added, then considering level 2 stations is important, provided that the location of these additional vehicles is planned in advance.

7.5.4 ALTERNATIVE OPERATING SCENARIOS

Besides informing the GPCG about current planning practices – seasonal vehicles and fleet expansion – it is also interesting to study alternative scenarios to gain insights about possible managerial decisions which may improve the system's performance. For this purpose, four alternative scenarios are tested.

Firstly, the effect of removing legislation constraints concerning minimum vehicles at hospitals is assessed by lifting constraints (5) from the base line scenario. Secondly, a free vehicle relocation scenario is analysed, in which it is tested if the current system could be improved using only existing vehicles and existing stations but improving the way these vehicles are dynamically distributed among stations. For this purpose, a set of experiments are conducted allowing vehicles to be positioned at any station and considering all vehicles as selectable. Thirdly, a green-field scenario is used to test what would be an ideal system if the current system was not in place. To test this scenario, the constraints which set the initial state of the system (10-11) are removed. This way, the model can choose to open any station at the beginning of the planning horizon. However, the distinction between selectable and non-selectable vehicles and stations is still used. Finally, a fully flexible system is considered. In this case, besides assuming a green-field scenario, vehicle and station relocation constraints are lifted, and vehicles can be positioned at any station. Therefore, vehicles can be repositioned throughout the year to whichever stations maximize expected coverage. Although, in practice, this scenario may be excessively challenging to be implemented, it can provide insights on the maximum performance the system can achieve with the same resources – vehicles and stations – currently in operation.

Given the limited time frame of this dissertation and the fact that the second and third runs of the lexicographic method proved to be computationally challenging for most scenarios, only the first objective is considered hereafter. Further investigation of the remaining objectives can be conducted for the most promising scenarios. The results are presented in Figure 31.

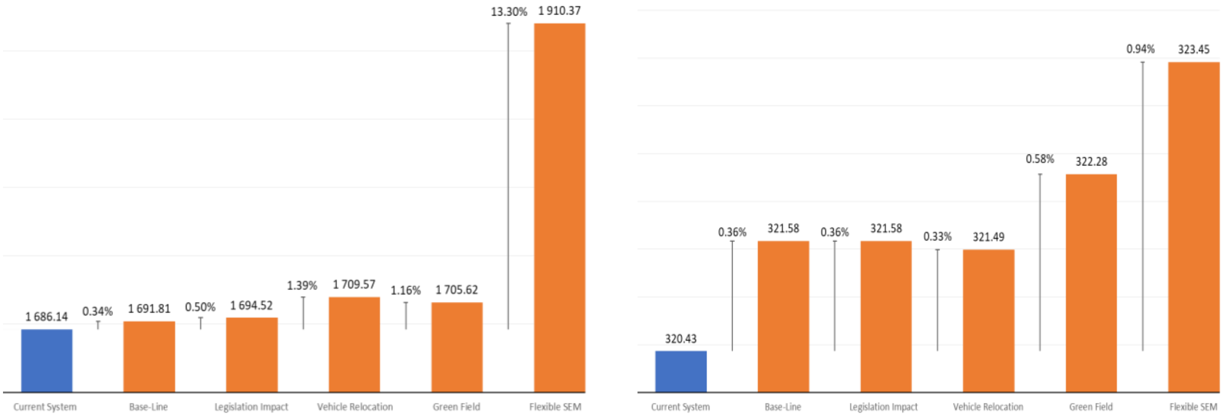


Figure 31 - Comparison of the results from the alternative operating scenarios (Left: Lisbon; Right: Setúbal).

Concerning the first experiment, the results suggest that the impact of legislative restrictions in isolation is small, mainly because the vehicles which are regulated by legislation – VMERs – are crewed by doctors from NHS hospitals and, as such, can only be located at the very sights in which the legislation requires them to be. The expected improvement in coverage for Lisbon is just 0.50%, due to the allocation of VMERs to *INEM – R. Infante D. Pedro*. In Setúbal, the results are exactly the same as the base-line optimization (improvement of 0.36%), given that *Hospital de Setúbal* is the only station which can be assigned VMERs.

Moreover, the results of the vehicle relocation scenario suggest that the performance on both regions could be improved with the current vehicles and stations (i.e. without seasonal stations nor vehicles). In Lisbon, an improvement of 1.39% is expected, while for Setúbal it is just 0.33%. Note that these relocations do not need to be permanent, that is, the vehicle can be repositioned at the beginning of the day but still be stationed at its original station for replenishment or parking during the night.

In order to assess the volatility of the solution proposed by this scenario, the number of relocations of vehicles of each type on each month is displayed in Figure 32.

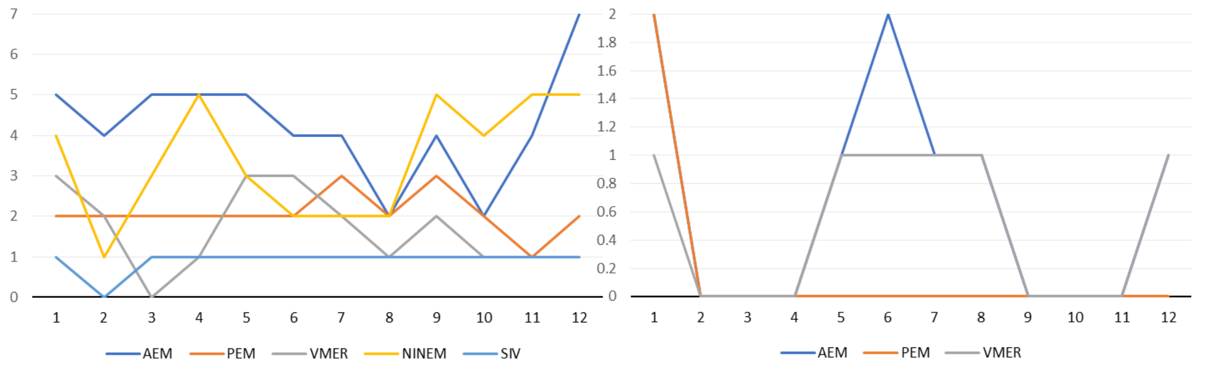


Figure 32 - Required relocations per month on the free vehicle relocation scenario (Left: Lisbon; Right: Setúbal).

The results suggest that, in Lisbon, an average of 12.3 relocations per month are required (47% of the fleet) while for Setúbal only 1.3 relocations are necessary (26% of the fleet). Given their higher flexibility, AEMs are the vehicles which are relocated more often. The SIV ambulance of Lisbon is also relocated on almost every month (with exception of February), while all PEMs are relocated on every month except November. In Setúbal, the PEMs are only relocated on the first month, while VMERs and AEMs are relocated at the beginning of the planning horizon as well as from May from August and in December.

Although it may be impossible to implement a system with a large amount of relocations, the results do highlight that performance improvements are possible without establishing new cooperation protocols for facilities. However, these improvements are modest, especially if the high number of required relocations is considered. Therefore, all in all, this strategy seems to be inefficient.

Concerning the green-field scenario, only modest improvements are attained once again: 1.16% for Lisbon and 0.58% for Setúbal. The resulting emergency stations are presented in Figure 33, including seasonal stations. This figure includes both seasonal and regular stations, as well as the comparison between existing and proposed stations.

Analysing Figure 33, it is interesting to note that the proposed distribution of emergency stations in both Lisbon and Setúbal is more scattered than the original solution. Note that, in Lisbon, the number of emergency stations in the north and west of the city is clearly increased. Nevertheless, 11 of the original stations are kept, while nine are permanently closed. Out of the 11 stations which are kept, three are necessary, by law (since they must house VMERs), while three others are the ones which can house NINEMs. In Setúbal, four stations are kept – *Hospital de Setúbal*, *BV Setúbal – Sede, Cruz Vermelha* and *BV Setúbal - Azeitão*, while the remaining

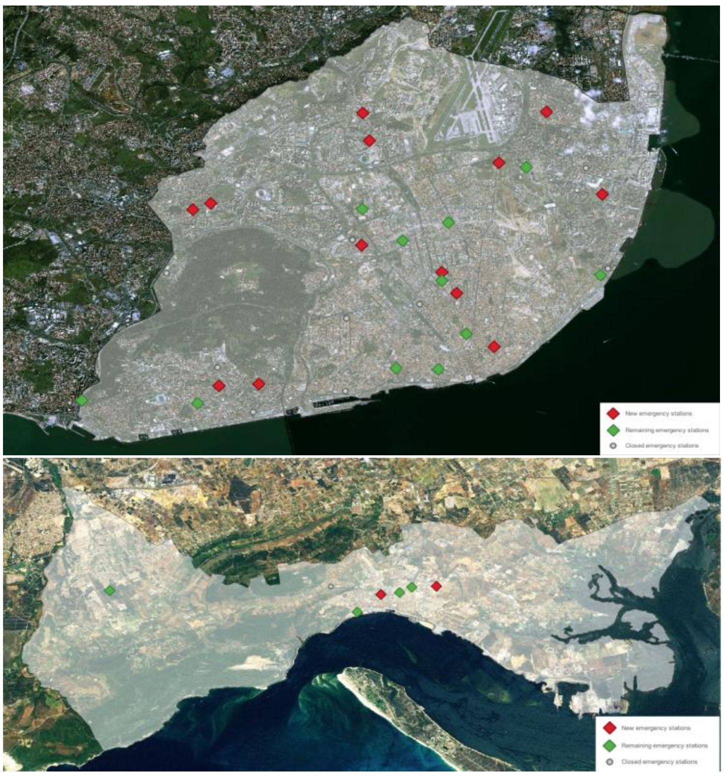


Figure 33 - Station configuration for the green-field scenario.

station is replaced by *Centro de Saúde Bonfim.* An additional seasonal station (July, August and September) is opened in *Centro de Saúde São Sebastião*.

Regarding the distribution of emergency vehicles in this scenario, in Lisbon VMERs are kept at their original sites (*Hospital São Francisco Xavier, Santa Maria* and *São José*) due to legislation. Similarly, NINEMs are assigned to firefighter corporations – *BV Beato, BV Campo de Ourique, RSB – Companhia de Intervenção Especial, Delegação de Lisboa da CVP*, *BV Lisboa* and *Corpo de Bombeiros Municipais – Quartel nº 2*, which receive both a NINEM and a PEM. AEMs, on the other hand, are relocated as the demand pattern changes. For instance, in the first period AEMs are mainly distributed by existing stations and health centers, such as: a) *Centros de Saúde Olivais, Graça, Benfica, Lumiar, Alcântara*, *Lóios – Olivais;* b) hospitals, *as Hospital da Cruz Vermelha, Santa Maria, São José, Pulido Valente*, *Instituto Português de Oncologia*; c) as well as police stations (*GNR Cavalaria Ajuda*) and both INEM facilities (*INEM Sede* and *INEM – R. Infante D. Pedro*.

In Setúbal, PEMs and VMERs are kept at their stations due to legislation and the fact that there are no other stations for PEMs. AEMs are positioned at *Hospital de Setúbal* and *Centro de Saúde Bonfim*, being repositioned from *Hospital de Setúbal* to *Centro de Saúde de São Sebastião* during the summer.

Finally, the fully flexible EMS is, as expected, the scenario which yields a more significant improvement. For Lisbon, using the same vehicles and number of emergency stations, an improvement of 13.1% can be achieved. This improvement is higher than what would be obtained by adding four SIVs to the existing fleet. In Setúbal, a 0.94% improvement is possible which, in turn, is less attractive than the addition of a single SIV, which leads to improvements of 2.66%. Note, however, that for both cities, the improvement in a flexible system – in which both stations and vehicles are allowed to be relocated – is 3 to 10 times higher than allowing only vehicles to be relocated or scratching off the existing stations. Therefore, it can be concluded that allowing multiple vehicle relocations without also providing flexibility for the location of emergency stations is ineffective.

7.6 GENERAL RECOMMENDATIONS

From the previous discussion, it can be concluded that:

- Complying with the current restrictions imposed by INEM, the legislation and SIEM partners do not enable great improvements to be attained in what concerns expected coverage, mainly because the configuration of existing stations cannot be changed, and many vehicles cannot be relocated;
- Seasonal PEMs improve the system performance in both areas, but would be more effective if deployed during the winter months for Lisbon;
- If the fleet is to be expanded, the addition of SIVs to Setúbal is the most promising course of action, followed by AEMs. Further PEMs or VMERs are not beneficial due to the lack of available stations. For Lisbon, both AEMs and SIVs are good candidates, while the benefit of additional VMERs is far more limited. Accounting for the purchasing cost of vehicles is paramount to develop a cost-benefit analysis and assess the impact per unit cost of each vehicle type;

- Furthermore, accounting for Level 2 stations for potential expansion scenarios does not appear to bring considerable benefits compared to level 1 stations for small fleet expansions. However, if larger expansions are predicted, than their impact may be beneficial;
- Only if both vehicles and stations are allowed to be relocated can the system be substantially improved;
- With the current resources, the maximum expected coverage improvement for Lisbon is 13.3% increase and, for Setúbal, 0.98%.

7.7 SENSITIVITY ANALYSIS

Given the inherent uncertainty of several model parameters as well as the need to estimate some inputs, a sensitivity analysis is carried out to assess how the solutions provided by the model are likely to change due to variations in input parameters. For this purpose, a ±10% variation (for continuous parameters) and ±1 (for discrete parameters) is introduced in those parameters subject to uncertainty (such as demand) or those which had to be estimated (including capacity, operating and assignment costs, travel and service time, and coverage probabilities). Again, a lexicographic approach is used for the second and third objectives. Figure 34 shows the results of this analysis, which are conducted on the base line instance for Setúbal.

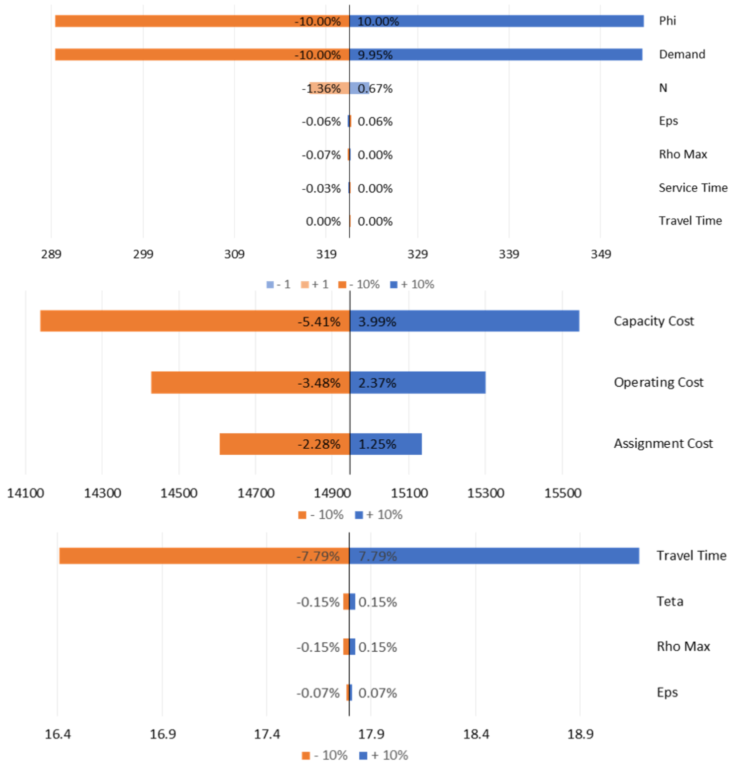


Figure 34 - Sensitivity Analysis results (Top: coverage; Middle: cost; Bottom: equity).

Concerning the first objective, the parameters which have a larger impact are demand and coverage probabilities. This result could be expected, since these parameters are directly used in the computation of this objective function, which has only one term. The value of N also has a considerable impact of the resulting objective since it influences the allocation of emergency requests to vehicles, promoting more distributed solutions but also leading to allocations which are less efficient. Surprisingly, the value of

ρ^{max} (the maximum vehicle workload) has little impact in coverage, although it is important to highlight that the actual variation of this input is small and greater variations could be applied. The remaining parameters – Travel Time, Service Time and ε –show little influence in the value of expected coverage.

Regarding cost, the cost category which generates a higher effect in total cost is capacity cost, followed by the operating cost and, finally, assignment cost. Note that the cost categories with higher impact correspond to those which are less flexible. That is, capacity costs are paid whenever a vehicle is allocated to a station, whereas operating costs are incurred only if that vehicle is operating and, finally, assignment costs are only paid if the vehicle is dispatched. As such, there are less degrees of freedom to adjust capacity costs than operating or assignment costs, which may help explain why the first category is the one with highest impact.

Finally, concerning the equity objective, not surprisingly, travel time has a significant effect on the objective function value. The value of θ (maximum travel time) presents a smaller variation because, in the studied instance, most demand is assigned to vehicles, thus very few requests are penalized at the maximum system travel time. Again, ρ^{max} and ε do not appear to significantly influence this objective.

7.8 HEURISTIC APPROACH RESULTS

Finally, the performance of the proposed two-stage hybrid heuristic is compared to CPLEX. For this purpose, the instances used in the scalability analysis are applied. Since the heuristic is only applicable to the first objective, only the first run (run 0) is considered. Recall that the heuristic requires two parameters (N and ε) which control the addition of cuts to SP1. In order to derive the value of these parameters, a grid search is conducted trying several parameter combinations and letting the heuristic run for 10 iterations. The results for the L1.0 instance are presented in Table 21.

Table 21 - Grid search results for the parameters of the heuristic.

	N						
ε	**20**	**30**	**35**	**39**	**40**	**50**	**60**
1	1692.664	1692.103	1691.841	1691.800	1691.990	1690.204	1690.204
1.25	1692.848	1693.196	1691.840	1691.800	1691.990	1690.204	1690.204
1.5	1692.824	1693.196	1693.182	1691.800	1691.990	1690.204	1690.204
1.75	1691.715	1692.988	1693.021	**1693.194**	1691.990	1690.204	1690.204
2	1692.084	1692.667	1693.138	1693.163	1693.235	1690.204	1690.204
2.5	1691.706	1692.541	1692.782	1692.848	1692.855	1690.204	1690.204
3	1691.855	1692.089	1692.582	1692.629	1692.853	1690.204	1690.204

The grid search results show that, after a certain value of N, the result of the heuristic remains unchanged regardless of the selected value for ε. Furthermore, greater values of N generally require also higher values of ε to obtain improved results.

Having selected the most promising heuristic parameters, Table 22 presents the results, comparing the CPLEX statistics with a simple problem decomposition (just one iteration of the heuristic), two iterations and five iterations.

Table 22 - Results of the hybrid heuristic in comparison with CPLEX.

	CPLEX				Problem Decomposition			
Experiment	**Objective**	**Gap**[4]	**Time**	**Best-Bound**	**Objective**	**Gap**	**Time**	**Δ**
L0.0	1691.43	0.47%	237.28	1699.38	1691.28	0.48%	92.41	-61.05%
L1.0	1648.13	0.40%	2951.37	1654.72	1649.67	0.31%	184.91	-93.73%
L2.0	1652.12	0.18%	9191.09	1655.09	1650.06	0.30%	410.07	-95.54%
S0.0	323.39	0.49%	22.44	324.978	317.71	2.29%	17.04	-24.08%
S1.0	322.41	0.00%	51.19	322.406	315.79	2.09%	34.83	-31.96%
S2.0	321.58	0.00%	247.64	321.584	314.29	2.32%	75.65	-69.45%

	Heuristic (2 iterations)				Heuristic (5 iterations)			
Experiment	**Objective**	**Gap**	**Time**	**Δ**	**Objective**	**Gap**	**Time**	**Δ**
L0.0	1693.14	0.37%	144.64	-39.04%	1693.14	0.37%	307.04	+29.40%
L1.0	1650.75	0.24%	295.59	-89.98%	1650.75	0.24%	612.56	-79.25%
L2.0	1650.85	0.26%	654.91	-92.87%	1650.85	0.26%	1432.67	-84.41%
S0.0	319.55	1.70%	24.33	+8.42%	320.51	1.39%	43.80	+95.19%
S1.0	316.39	1.90%	46.26	-9.62%	319.14	1.02%	81.29	+58.79%
S2.0	314.29	2.32%	104.17	-57.93%	315.03	2.08%	193.13	-22.01%

The results show that the heuristic provides good feasible solutions for both small and large test instances. In fact, the relative gap of the solutions yielded by the heuristic is never greater than 2.10% after five iterations. Even with just one iteration (solving the two subproblems in sequence), the solutions are relatively good, especially considering that the computational times are reduced by up to 95.54% from CPLEX. When further iterations are allowed, the solution is further improved, but only slightly. This may suggest that, although the proposed cuts do improve the solutions, more effective cuts may exist.

Nevertheless, for small instances (namely, S0.0 and S1.0), the gains in computational time are small for few iterations. For five iterations, the heuristic takes more time than CPLEX and produces worse results. Therefore, it can be concluded that the heuristic should not be applied for small instances, for which CPLEX is clearly more efficient and effective. On the other hand, for larger instances, the computational gains of the heuristic are attractive. For both the L1.0 and L2.0 instances, the heuristic outperforms CPLEX after just two iterations, reducing the computational time by 39.04% and 89.98%, respectively. Additionally, in the larger test instance (L2.0), the heuristic yields a slightly worse solution (a gap of 0.26% compared to 0.18% of CPLEX) but the computational time is reduced by 84.41%.

It is also interesting to highlight that, after a given number of iterations (most often less than 10), the heuristic procedure stalls. That is, after a significant number of cuts are added to SP1, the new cuts become redundant and the heuristic becomes stuck in the same solution. This phase is reached more quickly if N is larger, given that more cuts are added at the beginning of the algorithm. Although this property may help reduce the computational burden of the heuristic (since it can be stopped after two consecutive iterations with the same solution), it could be interesting to explore strategies to diversify the search and enable the algorithm to search other regions of the search space.

It is possible to conclude that the proposed heuristic is an effective solution procedure for large instances, but not appropriate for smaller instances. Additionally, it suffers from the major drawback that it is designed to handle a single objective.

[4] The relative MIP gap is calculated with the same formula as CPLEX, to enable the comparison of results: $Gap\ (\%) = \frac{|Best-Bound-Integer\ Solution|}{10^{-10}+\ |Integer\ Solution|}$ (IBM Knowledge Center, 2019b).

8. CONCLUSIONS AND FUTURE WORK

EMS are highly complex systems which are designed to save lives, being paramount in improving the health outcomes of the population. To do so, EMS systems usually operate multiple tiers of emergency vehicles, interacting with dispatching centres and NHS hospitals, while coordinating many entities which cooperate to provide the best possible care.

In this dissertation, the Portuguese EMS system (SIEM), managed by INEM, is studied. The SIEM has evolved since the 1960's, and nowadays answers more than 1.300.000 calls per year. Its main features include deploying several types of vehicles operated by multiple public entities with varying levels of responsibility, providing two layers of medical care (BLS and ALS) and using existing facilities as stations for emergency vehicles. Despite having to manage such a challenging system, INEM planners must rely on experience and intuition to make complex planning decisions. One of these decisions consists of determining the long-term positions of vehicles while waiting for emergency requests. This is both a strategical and tactical problem, as it encompasses the selection of stations and the allocation of vehicles to these stations. Naturally, this decision has a profound impact on the system's response which, in turn, greatly affects the medical outcomes of patients. Despite the importance of this decision and the multiple legal and self-imposed constraints, INEM's current planning practices are still mostly guided by empirical methods. This may help explain why SIEM's performance is below the international standards, although the system includes more emergency vehicles than many countries.

Recognizing the potential of a more sophisticated, scientific approach, the primary goal of this research is to apply OR techniques to assist INEM's vehicle planning processes. Additional goals include characterizing EMS systems in general; providing an in-depth review of the wide literature of optimization models for EMS planning and contributing to the literature by modelling, formulating and implementing an optimization model in a real EMS context.

To support the attainment of these goals, an exhaustive literature review of facility location models in the EMS context is provided, organized around modelling approaches. It is concluded that this problem has been deeply studied under multiple approaches, ranging from simple to very complex models capturing different aspects of EMS systems. Despite the richness of this research stream, very few studies incorporate multiple EMS aspects within a simple and integrated approach. Additionally, most approaches consider a green-field scenario, thus ignoring the fact that the current system is in operation and that, in practice, the main objective should be to help planners redesign existing systems.

Exploring the insights of the literature review and the features of the case-study, a Multi-Objective Dynamic MIP model is proposed. Its main goal is to assist the gradual reconfiguration of an existing EMS system over time by determining the configuration of emergency stations on different periods, the allocation of emergency vehicles among these stations, and the areas of responsibility of each vehicle. The model considers (1) several types of emergency requests with varying needs (care levels), (2) multiple vehicles with different capabilities which may be dispatched together to a call (multi-dispatch), (3) strategic and tactical vehicle and station relocations, (4) fluctuations in demand and travel times both during the day and throughout a wider planning horizon. Furthermore, three objectives are considered: expected coverage, cost and equity of access.

Besides, given the known complexity of solving large MIP models, a hybrid heuristic approach based on decomposing the problem into two sub-problems is proposed. This heuristic is intended to streamline model solution for the first objective – expected coverage – enabling, for instance, the integration of the model in a potential decision-support system.

In order to apply the model to the case-study, two historically challenging regions are considered: Lisbon and Setúbal. The analysed planning horizon is the year of 2020 (12 months). The data collection and treatment procedures necessary to transform real data into model inputs are described.

A Lexicographic approach is proposed to handle multiple objectives, given the hierarchical structure of the objectives for the decision-maker: expected coverage is the most critical objective, followed by cost and equity. The results suggest that, under the current restrictions, only slight improvements are possible. In fact, comparing with the estimated current system performance, only a modest improvement of 0.31% and 0.36% in expected coverage for Lisbon and Setúbal, respectively, are obtained. Although for the remaining objectives more attractive improvements can be obtained, these results show that INEM restricts the opening and closing of facilities and relocation of vehicles to a great extent, thus only enabling very slight adjustments. Therefore, unless INEM is willing to able some of its current self-imposed restrictions, there is only a small room for improvement.

For this reason, several alternative policy scenarios are studied. In a first phase, current planning practices are examined, namely the addition of seasonal PEMs during the summer. The results show that the seasonal vehicles do improve the system performance, even if only slightly. The suggested positions for seasonal PEMs are *BV Setúbal – Azeitão* for Setúbal and *BV Cabo Ruivo* in Lisbon. Furthermore, in what concerns the deployment months for the seasonal PEMs, it is concluded that Lisbon would benefit more if such vehicles were deployed during the winter months, while for Setúbal the current deployment months seem appropriate.

The possibility of expanding the fleet size with the addition of new vehicles to the existing fleet is also studied. In Lisbon, the addition of one AEM is the most promising course of action, leading to improvements of around 3.1%. However, if further vehicles are to be added, then the addition of SIVs, with their greater flexibility, is recommended. Similarly, for Setúbal, the addition of SIVs is also recommended. For instance, four SIVs can improve Lisbon's performance by 10.5%, while for Setúbal the improvement is 3.09%. Tests also show that considering a broader set of facilities as potential base stations (schools and police facilities) besides health units does not provide a significant benefit unless a large expansion is planned for in advance.

Furthermore, four scenarios are studied: the impact of legislation, a free vehicle relocation system, a green-field scenario and a fully flexible EMS. The results from these scenarios enable the comparison of the expected system performance under increasing levels of flexibility. The main conclusions are that the impact of legislation is small, and that performance can be greatly improved with existing resources provided that both stations and vehicles are allowed to be relocated.

All in all, the model proves to be quite flexible in the scenarios it enables to consider, clearly presenting the decision-maker how the system can and should evolve for varying degrees of flexibility. Besides the insights for INEM, the main contributions of this research are (1) a model of a EMS system with multiple vehicles and call priorities, (2) considering of the existing system in planning decisions, (3) integration of the three main concerns of EMS planners which had not been combined and (4) a hybrid heuristic capable of reducing the computational burden for large instances concerning the coverage objective.

Given the limitations in time and data availability, several assumptions are required to estimate parameters for the model. Namely, travel times and coverage probabilities are estimated from public information by applying an empirical model from the literature; Euclidean distances are used for demand-aggregation; travel times and coverage probabilities are assumed to be independent of the vehicle and to remain unchanged throughout the planning horizon and station capacities are also assumed to be constant. Moreover, cost information could not be provided by INEM, so rough estimates are presented based on public information. Therefore, in the future, more information should be collected concerning opening and closing costs, both from the GPCG as well as other INEM departments and, if necessary, SIEM partners. Once more cost information is available, a cost-benefit analysis is also suggested for the fleet expansion scenarios.

Besides overcoming limitations regarding data collection and treatment, further suggestions for future work can also be identified. Regarding the mathematical model, it could be interesting to explore alternative methods to determine the upper bound on vehicle workload and the minimum number of vehicles which must be considered responsible for a given call. In fact, the proposed approach based on the Erlang-C formula implies assuming that vehicles operate as independent servers and that service time is exponentially distributed, both of which are limiting assumptions. One alternative could be to develop and apply a Hypercube model which fits the features of the model: multiple customer classes, server types with multi-dispatch to the same call and heterogeneous service times. Another alternative would be the use of an iterative simulation-optimization approach. Further research is also necessary to evaluate the impact of the chosen reliability level and other vehicle availability parameters in the resulting model solutions. Moreover, the proposed model considers deterministic demand. In the future, it may be appropriate to include demand stochasticity via, for instance scenarios or, alternatively, chance constraints. However, these modifications are expected to increase the model's complexity and, simultaneously, its computational burden. Also, alternative equity measures may be pursued, since the proposed alternative only considers one dimension of equity. Another potential improvement consists of accounting for the impact of external vehicles (and external emergency requests) in the performance of the model. The current model assumes that the region of interest is isolated from its surroundings. However, in reality, there are adjacent areas whose demand and vehicles impact the operation of both regions. This can be modelled, for instance, as an additional demand node. Additionally, once precise cost information is collected, it may be appropriate to rethink the cost categories included in the model.

Concerning the solution approach, the heuristic can be improved by seeking diversification strategies and alternative cuts to polish the solution. Another suggestion is the development of heuristics designed to handle the remaining objectives, allowing for study of cost and equity concerns in the alternative scenarios. It can also be interesting to study alternative hybrid heuristics, exploring other decomposition possibilities: solving sequentially for shorter planning horizons, dividing the region of interest into sub-areas or splitting different emergency categories.

Finally, following the literature (see Chapter 4), it is suggested that the model is validated against a simulation model, to identify opportunities for improvement.

BIBLIOGRAPHY

Al-Shaqsi, S. (2010) 'Models of International Emergency Medical Service (EMS) Systems', *Oman Medical Journal*, 25(4), pp. 320–324. doi: 10.5001/omj.2010.92.

Alanis, R., Ingolfsson, A. and Kolfal, B. (2013) 'A markov chain model for an EMS system with repositioning', *Production and Operations Management*. doi: 10.1111/j.1937-5956.2012.01362.x.

Alsalloum, O. I. and Rand, G. K. (2006) 'Extensions to emergency vehicle location models', *Computers and Operations Research*. doi: 10.1016/j.cor.2005.02.025.

Aly, A. A. and White, J. A. (1978) 'Probabilistic formulation of the emergency service location problem', *Journal of the Operational Research Society*. doi: 10.1057/jors.1978.261.

Amiri, A. (1998) 'The design of service systems with queueing time cost, workload capacities and backup service', *European Journal of Operational Research*. doi: 10.1016/S0377-2217(97)82111-X.

Amiri, A. (2001) 'Multi-hour service system design problem', *European Journal of Operational Research*. doi: 10.1016/S0377-2217(99)00397-5.

Ansari, S., McLay, L. A. and Mayorga, M. E. (2015) 'A Maximum Expected Covering Problem for District Design', *Transportation Science*. doi: 10.1287/trsc.2015.0610.

Aringhieri, R. *et al.* (2017) 'Emergency medical services and beyond: Addressing new challenges through a wide literature review', *Computers and Operations Research*. Elsevier, 78(1), pp. 349–368. doi: 10.1016/j.cor.2016.09.016.

Aringhieri, R., Carello, G. and Morale, D. (2016) 'Supporting decision making to improve the performance of an Italian Emergency Medical Service', *Annals of Operations Research*. doi: 10.1007/s10479-013-1487-0.

Associação Bombeiros para Sempre (2015) *Legislação: Valores Pagos Pelo INEM Aos Seus Parceiros*.

Baker, J. R. and Fitzpatrick, K. E. (1986) 'Determination of an optimal forecast model for ambulance demand using goal programming', *Journal of the Operational Research Society*. doi: 10.1057/jors.1986.182.

Başar, A., Çatay, B. and Ünlüyurt, T. (2011) 'A multi-period double coverage approach for locating the emergency medical service stations in Istanbul', *Journal of the Operational Research Society*. doi: 10.1057/jors.2010.5.

Başar, A., Çatay, B. and Ünlüyurt, T. (2012) 'A taxonomy for emergency service station location problem', *Optimization Letters*. doi: 10.1007/s11590-011-0376-1.

Batta, R., Dolan, June M. and Krishnamurthy, N. N. (1989) 'The Maximal Expected Covering Location Problem: Revisited', *Transportation Science*. doi: 10.1287/trsc.23.4.277.

Batta, R., Dolan, June M and Krishnamurthy, N. N. (1989) 'The Maximal Expected Covering Location Problem: Revisited', *Transportation Science*. INFORMS, 23(4), pp. 277–287. doi: 10.1287/trsc.23.4.277.

Batta, R. and Mannur, N. R. (1990) 'Covering-Location Models for Emergency Situations That Require Multiple Response Units', *Management Science*. doi: 10.1287/mnsc.36.1.16.

Bélanger, V., Ruiz, A. and Soriano, P. (2018) *Recent optimization models and trends in location, relocation, and dispatching of emergency medical vehicles*, *European Journal of Operational Research*. doi: 10.1016/j.ejor.2018.02.055.

Beraldi, P. and Bruni, M. E. (2009) 'A probabilistic model applied to emergency service vehicle location', *European Journal of Operational Research*. doi: 10.1016/j.ejor.2008.02.027.

Beraldi, P., Bruni, M. E. and Conforti, D. (2004) 'Designing robust emergency medical service via stochastic programming', *European Journal of Operational Research*. doi: 10.1016/S0377-2217(03)00351-5.

Van Den Berg, P. L. and Aardal, K. (2015) 'Time-dependent MEXCLP with start-up and relocation cost', *European Journal of Operational Research*. Elsevier Ltd., 242(2), pp. 383–389. doi: 10.1016/j.ejor.2014.10.013.

Van Den Berg, P. L., Kommer, G. J. and Zuzáková, B. (2016) 'Linear formulation for the Maximum Expected Coverage Location Model with fractional coverage', *Operations Research for Health Care*. Elsevier Ltd, 8, pp. 33–41. doi: 10.1016/j.orhc.2015.08.001.

Van Den Berg, P. L., Legemaate, G. A. G. and Van Der Mei, R. D. (2017) 'Increasing the responsiveness of firefighter services by relocating base stations in Amsterdam', *Interfaces*. doi: 10.1287/inte.2017.0897.

Berman, O. *et al.* (2009) 'The variable radius covering problem', *European Journal of Operational*

Research. doi: 10.1016/j.ejor.2008.03.046.

Berman, O., Drezner, Z. and Krass, D. (2010a) 'Cooperative cover location problems: The planar case', *IIE Transactions (Institute of Industrial Engineers)*. doi: 10.1080/07408170903394355.

Berman, O., Drezner, Z. and Krass, D. (2010b) 'Generalized coverage: New developments in covering location models', *Computers and Operations Research*. doi: 10.1016/j.cor.2009.11.003.

Berman, O., Hajizadeh, I. and Krass, D. (2013) 'The maximum covering problem with travel time uncertainty', *IIE Transactions (Institute of Industrial Engineers)*. doi: 10.1080/0740817X.2012.689121.

Berman, O. and Krass, D. (2002) 'The generalized maximal covering location problem', *Computers and Operations Research*. doi: 10.1016/S0305-0548(01)00079-X.

Berman, O., Krass, D. and Wang, J. (2011) 'The probabilistic gradual covering location problem on a network with discrete random demand weights', *Computers and Operations Research*. doi: 10.1016/j.cor.2011.01.005.

Berman, O. and Wang, J. (2011) 'The minmax regret gradual covering location problem on a network with incomplete information of demand weights', *European Journal of Operational Research*. doi: 10.1016/j.ejor.2010.08.016.

Bertsimas, D. and Ng, Y. (2019) 'Robust and stochastic formulations for ambulance deployment and dispatch', *European Journal of Operational Research*. doi: 10.1016/j.ejor.2019.05.011.

Bianchi, G. and Church, R. L. (1988) 'A hybrid fleet model for emergency medical service system design', *Social Science and Medicine*. doi: 10.1016/0277-9536(88)90055-X.

Borras, F. and Pastor, J. T. (2002) 'The Ex-Post Evaluation of the Minimum Local Reliability Level: An Enhanced Probabilistic Location Set Covering Model', *Annals OR*, 111, pp. 51–74. doi: 10.1023/A:1020941400807.

Budge, S., Ingolfsson, A. and Erkut, E. (2009) 'Technical Note—Approximating Vehicle Dispatch Probabilities for Emergency Service Systems with Location-Specific Service Times and Multiple Units per Location', *Operations Research*. doi: 10.1287/opre.1080.0591.

Budge, S., Ingolfsson, A. and Zerom, D. (2010) 'Empirical Analysis of Ambulance Travel Times: The Case of Calgary Emergency Medical Services', *Management Science*. doi: 10.1287/mnsc.1090.1142.

Cardoso, T. *et al.* (2015) 'An integrated approach for planning a long-term care network with uncertainty, strategic policy and equity considerations', *European Journal of Operational Research*. doi: 10.1016/j.ejor.2015.05.074.

Channouf, N. *et al.* (2007) 'The application of forecasting techniques to modeling emergency medical system calls in Calgary, Alberta', *Health Care Management Science*. doi: 10.1007/s10729-006-9006-3.

Chanta, S. *et al.* (2011) 'The minimum p-envy location problem: a new model for equitable distribution of emergency resources', *IIE Transactions on Healthcare Systems Engineering*. doi: 10.1080/19488300.2011.609522.

Chanta, S., Mayorga, M. E. and McLay, L. A. (2014a) 'Improving emergency service in rural areas: a bi-objective covering location model for EMS systems', *Annals of Operations Research*. doi: 10.1007/s10479-011-0972-6.

Chanta, S., Mayorga, M. E. and McLay, L. A. (2014b) 'The minimum p-envy location problem with requirement on minimum survival rate', *Computers and Industrial Engineering*. doi: 10.1016/j.cie.2014.06.001.

Chapman, S. C. and White, J. A. (1974) 'Probabilistic formulations of emergency service facilities location problems', in *ORSA/TIMS Conference Proceedings*.

Charnes, A. and Storbeck, J. (1980) 'A goal programming model for the siting of multilevel EMS systems', *Socio-Economic Planning Sciences*. doi: 10.1016/0038-0121(80)90029-4.

Cheu, R. L., Lei, H. and Aldouri, R. (2010) 'Optimal assignment of emergency response service units with time-dependent service demand and travel time', *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations*. doi: 10.1080/15472450.2010.516232.

Chiyoshi, F., Iannoni, A. P. and Morabito, R. (2011) 'A tutorial on hypercube queueing models and some practical applications in Emergency Service Systems', *Pesquisa Operacional*. doi: 10.1590/s0101-74382011000200005.

Cho, S.-H. *et al.* (2014) 'Simultaneous Location of Trauma Centers and Helicopters for Emergency Medical Service Planning', *Operations Research*. doi: 10.1287/opre.2014.1287.

Chong, K. C., Henderson, S. G. and Lewis, M. E. (2016) 'The Vehicle Mix Decision in Emergency Medical Service Systems', *Manufacturing & Service Operations Management*. doi: 10.1287/msom.2015.0555.

Chu, F. *et al.* (2018) 'Distribution-Free Model for Ambulance Location Problem with Ambiguous Demand', *Journal of Advanced Transportation*, 2018, pp. 1–12. doi: 10.1155/2018/9364129.

Chuang, C. L. and Lin, R. H. (2007) 'A maximum expected covering model for an ambulance location problem', *Journal of the Chinese Institute of Industrial Engineers*. doi: 10.1080/10170660709509061.

Church, R. L. (2002) 'Geographical information systems and location science', *Computers and Operations Research*. doi: 10.1016/S0305-0548(99)00104-5.

Church, R. L. and Gerrard, R. A. (2003) 'The multi-level location set covering model', *Geographical Analysis*. doi: 10.1111/j.1538-4632.2003.tb01115.x.

Church, R. L. and Meadows, M. E. (1979) 'Location Modeling Utilizing Maximum Service Distance Criteria', *Geographical Analysis*, 11(4), pp. 358–373. doi: 10.1111/j.1538-4632.1979.tb00702.x.

Church, R. L. and Roberts, K. L. (1983) 'Generalized coverage models and public facility location', *Papers of the Regional Science Association*. doi: 10.1007/978-3-319-40631-2_21.

Church, R. and ReVelle, C. (1974) 'The Maximal Covering Location Problem', *Papers of the Regional Science Association*, 32(1), pp. 101–118. doi: 10.1007/BF01942293.

Colombo, F., Cordone, R. and Lulli, G. (2016) 'The multimode covering location problem', *Computers and Operations Research*. doi: 10.1016/j.cor.2015.09.003.

Coskun, N. and Erol, R. (2010) 'An optimization model for locating and sizing emergency medical service stations', *Journal of Medical Systems*. doi: 10.1007/s10916-008-9214-0.

Current, J. R. and Storbeck, J. E. (1988) 'Capacitated Covering Models', *Environment and Planning B: Planning and Design*. doi: 10.1068/b150153.

Daskin, M. S. (1983) 'A Maximum Expected Covering Location Model: Formulation, Properties and Heuristic Solution', *Transportation Science*, 17(1), pp. 48–70.

Daskin, M. S. (1987) 'Location, Dispatching AND Routing Models For Emergency Services With Stochastic Travel Times', in A. Ghosh, & G. R. (ed.) *Spatial analysis and location-allocation models*. New York, NY, pp. 224–265.

Daskin, M. S. *et al.* (1989) 'Aggregation effects in maximum covering models', *Annals of Operations Research*. doi: 10.1007/BF02097799.

Daskin, M. S. (1995) *Network and Discrete Location - Models, Algorithms, and Applications*. JOHN WILEY & SONS, INC.

Daskin, M. S. and Stern, E. H. (1981) 'A Hierarchical Objective Set Covering Model for Emergency Medical Service Vehicle Deployment', *Transportation Science*, 15(2), pp. 137–152. doi: 10.1287/trsc.15.2.137.

Davari, S., Fazel Zarandi, M. H. and Hemmati, A. (2011) 'Maximal covering location problem (MCLP) with fuzzy travel times', *Expert Systems with Applications*. doi: 10.1016/j.eswa.2011.05.031.

Davaria, S. *et al.* (2010) 'The variable radius covering problem with fuzzy travel times', in *2010 IEEE World Congress on Computational Intelligence, WCCI 2010*. doi: 10.1109/FUZZY.2010.5584133.

Davis, S. G. (1981) 'Analysis of the deployment of emergency medical services', *Omega*, 9(6), pp. 655–657. doi: https://doi.org/10.1016/0305-0483(81)90054-2.

Davoudpour, H., Mortaz, E. and Hosseinijou, S. A. (2014) 'A new probabilistic coverage model for ambulances deployment with hypercube queuing approach', *International Journal of Advanced Manufacturing Technology*. doi: 10.1007/s00170-013-5336-8.

Degel, D. *et al.* (2015) 'Time-dependent ambulance allocation considering data-driven empirically required coverage', *Health Care Management Science*. doi: 10.1007/s10729-014-9271-5.

Delignette-Muller, M.-L. *et al.* (2019) 'Package "fitdistrplus": Help to Fit of a Parametric Distribution to Non-Censored or Censored Data'. Available at: http://listes.univ-lyon1.fr/wws/info/fitdist-users.

Dibene, J. C. *et al.* (2017) 'Optimizing the location of ambulances in Tijuana, Mexico', *Computers in Biology and Medicine*. doi: 10.1016/j.compbiomed.2016.11.016.

Dick, W. F. (2003) 'Anglo-American vs. Franco-German Emergency Medical Services System', *Prehospital and Disaster Medicine*. Cambridge University Press, 18(01), pp. 29–37. doi: 10.1017/S1049023X00000650.

Doerner, K. F. *et al.* (2005) 'Heuristic solution of an extended double-coverage ambulance location problem for Austria', *Central European Journal of Operations Research*.

Drezner, T., Drezner, Z. and Goldstein, Z. (2010) 'A stochastic gradual cover location problem', *Naval Research Logistics*. doi: 10.1002/nav.20410.

Drezner, T., Drezner, Z. and Guyse, J. (2009) 'Equitable service by a facility: Minimizing the Gini coefficient', *Computers and Operations Research*. doi: 10.1016/j.cor.2009.02.019.

Drezner, Z., Marianov, V. and Wesolowsky, G. O. (2016) ‘Maximizing the minimum cover probability by emergency facilities’, *Annals of Operations Research*, 246(1–2), pp. 349–362. doi: 10.1007/s10479-014-1726-z.

Drezner, Z., Wesolowsky, G. O. and Drezner, T. (2004) ‘The gradual covering problem’, *Naval Research Logistics*. doi: 10.1002/nav.20030.

Eaton, D. J. *et al.* (1985) ‘Determining emergency medical deployment in Austin, Texas’, *Interfaces*.

EENA (2018) *Public Safety Answering Points Global Edition*.

Emir-Farinas, H. and Francis, R. L. (2005) ‘Demand point aggregation for planar covering location models’, in *Annals of Operations Research*. doi: 10.1007/s10479-005-2044-2.

Erkut, E. (2008) ‘Maximum availability/reliability models for selecting ambulance station and vehicle locations: A critique’.

Erkut, E. *et al.* (2009) ‘Computational comparison of five maximal covering models for locating ambulances’, *Geographical Analysis*. doi: 10.1111/j.1538-4632.2009.00747.x.

Erkut, E., Erdogan, G. and Ingolfsson, A. (2008) ‘Ambulance location for maximum survival’, *Naval Research Logistics*. doi: 10.1002/nav.20267.

Ester, M. *et al.* (1996) ‘A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise’, *Kdd*. doi: 10.1016/B978-044452701-1.00067-3.

Farahani, R. Z. *et al.* (2012) ‘Covering problems in facility location: A review’, *Computers and Industrial Engineering*. doi: 10.1016/j.cie.2011.08.020.

Felder, S. and Brinkmann, H. (2002) ‘Spatial allocation of emergency medical services: Minimising the death rate or providing equal access?’, *Regional Science and Urban Economics*. doi: 10.1016/S0166-0462(01)00074-6.

Fischer, M. *et al.* (2011) ‘Comparing emergency medical service systems-A project of the European Emergency Data (EED) Project’, *Resuscitation*. doi: 10.1016/j.resuscitation.2010.11.001.

Francis, R. L. *et al.* (2009) ‘Aggregation error for location models: Survey and analysis’, *Annals of Operations Research*. doi: 10.1007/s10479-008-0344-z.

Fu, M. C., Glover, F. W. and April, J. (2005) ‘Simulation optimization: A review, new developments, and applications’, in *Proceedings - Winter Simulation Conference*. doi: 10.1109/WSC.2005.1574242.

Galvão, R. D., Chiyoshi, F. Y. and Morabito, R. (2005) ‘Towards unified formulations and extensions of two classical probabilistic location models’, *Computers and Operations Research*. doi: 10.1016/S0305-0548(03)00200-4.

Gerardo, Á. (2017) *Triagem de pedidos de assistência médica*. Available at: https://repositorio.iscte-iul.pt/handle/10071/15100.

Goldberg, J. *et al.* (1990) ‘A simulation model for evaluating a set of emergency vehicle base locations: Development, validation, and usage’, *Socio-Economic Planning Sciences*. doi: 10.1016/0038-0121(90)90017-2.

Goldberg, J. B. (2004) ‘Operations Research Models for the Deployment of Emergency Services Vehicles’, *EMS Management Journal*, 1(1), pp. 20–39. Available at: academic.csuohio.edu/holcombj/Deployment.pdf.

Goldberg, J. and Paz, L. (1991) ‘Locating Emergency Vehicle Bases When Service Time Depends on Call Location’, *Transportation Science*. doi: 10.1287/trsc.25.4.264.

Gomes, E. *et al.* (2004) ‘International EMS systems: Portugal’, *Resuscitation*, 62(3), pp. 257–260. doi: 10.1016/j.resuscitation.2004.06.009.

Goodchild, M. F. (1979) ‘The Aggregation Problem in Location-Allocation’, *Geographical Analysis*. doi: 10.1111/j.1538-4632.1979.tb00692.x.

Grannan, B. C., Bastian, N. D. and McLay, L. A. (2014) ‘A maximum expected covering problem for locating and dispatching two classes of military medical evacuation air assets’, *Optimization Letters*. doi: 10.1007/s11590-014-0819-6.

Hall, W. K. (1971) ‘Management science approaches to the determination of urban ambulance requirements’, *Socio-Economic Planning Sciences*. doi: 10.1016/0038-0121(71)90007-3.

Harewood, S. I. (2002) ‘Emergency ambulance deployment in barbados: A multi-objective approach’, *Journal of the Operational Research Society*. doi: 10.1057/palgrave.jors.2601250.

Henderson, S. G. and Mason, A. J. (2006) ‘Ambulance Service Planning: Simulation and Data Visualisation’, in *Operations Research and Health Care*. doi: 10.1007/1-4020-8066-2_4.

Hogan, K. and ReVelle, C. (1986) ‘Concepts and Applications of Backup Coverage’, *Management Science*, 32(11), pp. 1434–1444. doi: 10.1287/mnsc.32.11.1434.

Iannoni, A. P., Morabito, R. and Saydam, C. (2009) 'An optimization approach for ambulance location and the districting of the response segments on highways', *European Journal of Operational Research*. doi: 10.1016/j.ejor.2008.02.003.

Iannoni, A. P., Morabito, R. and Saydam, C. (2011) 'Optimizing large-scale emergency medical system operations on highways using the hypercube queuing model', *Socio-Economic Planning Sciences*. doi: 10.1016/j.seps.2010.11.001.

IBM Corporation (2010) *Efficient modeling with the IBM ILOG CPLEX (White Paper)*.

IBM Knowledge Center (2019a) *IBM ILOG Script for OPL*, *ILOG CPLEX Optimization Studio 12.8.0*. Available at: https://www.ibm.com/support/knowledgecenter/SSSA5P_12.8.0/ilog.odms.ide.help/OPL_Studio/opllangref/topics/opl_langref_script.html (Accessed: 18 September 2019).

IBM Knowledge Center (2019b) *Relative MIP gap tolerance*, *List of CPLEX Parameters*. Available at: https://www.ibm.com/support/knowledgecenter/SSSA5P_12.7.1/ilog.odms.cplex.help/CPLEX/Parameters/topics/EpGap.html (Accessed: 4 October 2019).

Indriasari, V. *et al.* (2010) 'Maximal service area problem for optimal siting of emergency facilities', *International Journal of Geographical Information Science*. doi: 10.1080/13658810802549162.

INEM (2015) 'Guia Prático Transporte de doentes Entidades Isentas Alvará', pp. 1–13. Available at: http://www.inem.pt/wp-content/uploads/2017/06/01.pdf.

Ingolfsson, A. (2013) 'EMS Planning and Management', in, pp. 105–128. doi: 10.1007/978-1-4614-6507-2_6.

Ingolfsson, A., Budge, S. and Erkut, E. (2008) 'Optimal ambulance location with random delays and travel times', *Health Care Management Science*, 11(3), pp. 262–274. doi: 10.1007/s10729-007-9048-1.

Instituto Nacional de Emergência Médica (2013) 'Sistema Integrado de Emergência Médica', pp. 1–20.

Instituto Nacional de Emergência Médica (2015) *Plano Estratégico 2014 | 2016*.

Instituto Nacional de Emergência Médica (2016a) *Relatório Anual de Atividades e Contas*. Available at: https://www.inem.pt/wp-content/uploads/2017/09/Relatorio-de-Atividades-e-Contas-2016.pdf.

Instituto Nacional de Emergência Médica (2016b) *Relatório Anual de Meios de Emergência Médica*.

Instituto Nacional de Emergência Médica (2017a) *Plano Estratégico 2017 | 2019*.

Instituto Nacional de Emergência Médica (2017b) *Relatório Anual de Atividades e Contas*.

Instituto Nacional de Emergência Médica (2017c) *Relatório Anual Meios de Emergência Médica*.

Instituto Nacional de Emergência Médica (2018) *Relatório Anual: Integração VMER & SIV 2017*. doi: 10.1038/jid.2008.321.

Jagtenberg, C. (2016) *EFFICIENCY AND FAIRNESS IN AMBULANCE PLANNING*.

Jang, H. and Lee, T. (2015) 'Demand point aggregation method for covering problems with gradual coverage', *Computers and Operations Research*. doi: 10.1016/j.cor.2015.01.006.

Jarvis, J. P. (1985) 'Approximating the Equilibrium Behavior of Multi-Server Loss Systems', *Management Science*. doi: 10.1287/mnsc.31.2.235.

Jayaraman, V. and Srivastava, R. (1995) 'A service logistics model for simultaneous siting of facilities and multiple levels of equipment', *Computers and Operations Research*. doi: 10.1016/0305-0548(94)E0013-W.

Kanoun, I., Chabchoub, H. and Aouni, B. (2010) 'Goal Programming Model for Fire and Emergency Service Facilities Site Selection', *INFOR: Information Systems and Operational Research*. doi: 10.3138/infor.48.3.143.

Knight, V. A., Harper, P. R. and Smith, L. (2012) 'Ambulance allocation for maximal survival with heterogeneous outcome measures', *Omega*. doi: 10.1016/j.omega.2012.02.003.

Kolesar, P., Walker, W. and Hausner, J. (1975) 'Determining the Relation between Fire Engine Travel Times and Travel Distances in New York City', *Operations Research*, 23(4), pp. 614–627. doi: 10.1287/opre.23.4.614.

Krafft, T. *et al.* (2003) 'European Emergency Data Project (EED Project): EMS data-based Health Surveillance System', *The European Journal of Public Health*. doi: 10.1093/eurpub/13.suppl_1.85.

Kvalseth, T. O. and Deems, J. M. (1979) 'Statistical models of the demand for emergency medical services in an urban area', *American Journal of Public Health*. doi: 10.2105/AJPH.69.3.250.

Lahijanian, B., Zarandi, M. H. F. and Farahani, F. V. (2017) 'Double coverage ambulance location modeling using fuzzy traveling time', in *Annual Conference of the North American Fuzzy Information Processing Society - NAFIPS*. doi: 10.1109/NAFIPS.2016.7851626.

Laporte, G. and Nickel, S. (2015) 'Introduction to Location Science', in *Location Science*. doi:

10.1007/978-3-319-13111-5.

Larson, R. C. (1974) 'A hypercube queuing model for facility location and redistricting in urban emergency services', *Computers & Operations Research*. Pergamon, 1(1), pp. 67–95. doi: 10.1016/0305-0548(74)90076-8.

Larson, R. C. (1975) 'Approximating the Performance of Urban Emergency Service Systems', *Operations Research*. doi: 10.1287/opre.23.5.845.

Lei, H., Cheu, R. L. and Aldouri, R. (2010) 'Optimal Allocation of Emergency Response Service Units to Cover Critical Infrastructures with Time-Dependent Service Demand and Travel Time', *Transportation Research Record: Journal of the Transportation Research Board*. doi: 10.3141/2137-09.

Leknes, H. *et al.* (2017) 'Strategic ambulance location for heterogeneous regions', *European Journal of Operational Research*. doi: 10.1016/j.ejor.2016.12.020.

Lerner, E. B. *et al.* (2007) 'A Comprehensive Framework for Determining the Cost of an Emergency Medical Services System', *Annals of Emergency Medicine*. doi: 10.1016/j.annemergmed.2006.09.019.

Li, X. *et al.* (2011) 'Covering models and optimization techniques for emergency response facility location and planning: A review', *Mathematical Methods of Operations Research*. doi: 10.1007/s00186-011-0363-4.

Liu, Y. *et al.* (2016) 'A double standard model for allocating limited emergency medical service vehicle resources ensuring service reliability', *Transportation Research Part C: Emerging Technologies*. doi: 10.1016/j.trc.2016.05.023.

MacQueen, J. (1967) 'Some methods for classification and analysis of multivariate observations', in *Proceedings of the fifth Berkeley Symposium on Mathematical Statistics and Probability*.

Marianov, V. and ReVelle, C. (1994) 'The queuing probabilistic location set covering problem and some extensions', *Socio-Economic Planning Sciences*. doi: 10.1016/0038-0121(94)90003-5.

Marianov, V. and ReVelle, C. (1996) 'The queueing maximal availability location problem: A model for the siting of emergency vehicles', *European Journal of Operational Research*, 93(1), pp. 110–120. doi: 10.1016/0377-2217(95)00182-4.

Marianov, V. and Serra, D. (1998) 'Probabilistic, maximal covering location-allocation models from congested systems1', *Journal of Regional Science*, 38(3), pp. 401–424. doi: 10.1111/0022-4146.00100.

Matteson, D. S. *et al.* (2011) 'Forecasting emergency medical service call arrival rates', *Annals of Applied Statistics*. doi: 10.1214/10-AOAS442.

McConnel, C. E. and Wilson, R. W. (1998) 'The demand for prehospital emergency services in an aging society', *Social Science and Medicine*. doi: 10.1016/S0277-9536(97)10029-6.

McLay, L. A. (2009) 'A maximum expected covering location model with two types of servers', *IIE Transactions (Institute of Industrial Engineers)*. doi: 10.1080/07408170802702138.

McLay, L. A. and Mayorga, M. E. (2010) 'Evaluating emergency medical service performance measures', *Health Care Management Science*. doi: 10.1007/s10729-009-9115-x.

Melo, M. T., Nickel, S. and Saldanha-da-Gama, F. (2009) 'Facility location and supply chain management - A review', *European Journal of Operational Research*. doi: 10.1016/j.ejor.2008.05.007.

Melo, M. T., Nickel, S. and Saldanha Da Gama, F. S. (2006) 'Dynamic multi-commodity capacitated facility location: A mathematical modeling framework for strategic supply chain planning', *Computers and Operations Research*. doi: 10.1016/j.cor.2004.07.005.

Micheletti, A. *et al.* (2010) 'A stochastic model for simulation and forecasting of emergencies in the area of Milano', in *2010 IEEE Workshop on Health Care Management, WHCM 2010*. doi: 10.1109/WHCM.2010.5441259.

Murray, A. T. (2016) 'Maximal Coverage Location Problem: Impacts, Significance, and Evolution', *International Regional Science Review*, 39(1), pp. 5–27. doi: 10.1177/0160017615600222.

Myers, J. B. *et al.* (2008) 'Evidence-based performance measures for emergency medical services systems: A model for expanded EMS benchmarking', *Prehospital Emergency Care*. doi: 10.1080/10903120801903793.

NHS England (2017) *New ambulance standards*. Available at: https://www.england.nhs.uk/urgent-emergency-care/arp/ (Accessed: 7 March 2019).

Nickel, S., Reuter-Oppermann, M. and Saldanha-da-Gama, F. (2016) 'Ambulance location under stochastic demand: A sampling approach', *Operations Research for Health Care*. doi: 10.1016/j.orhc.2015.06.006.

Owen, S. H. and Daskin, M. S. (1998) 'Strategic facility location: A review', *European Journal of Operational Research*, 111(3), pp. 423–447. doi: 10.1016/S0377-2217(98)00186-6.

Pirkul, H. and Schilling, D. A. (1991) 'The Maximal Covering Location Problem with Capacities on Total Workload', *Management Science*. doi: 10.1287/mnsc.37.2.233.

R Core Team (2017) 'R: A Language and Environment for Statistical Computing'. Vienna, Austria: R Foundation for Statistical Computing. Available at: https://www.r-project.org/.

Rajagopalan, H. K., Saydam, C. and Xiao, J. (2008) 'A multiperiod set covering location model for dynamic redeployment of ambulances', *Computers and Operations Research*. doi: 10.1016/j.cor.2006.04.003.

Repede, J. F. and Bernardo, J. J. (1994) 'Developing and validating a decision support system for locating emergency medical vehicles in Louisville, Kentucky', *European Journal of Operational Research*. doi: 10.1016/0377-2217(94)90297-6.

Republica Portuguesa (2014) 'Portaria nº 260/2014 - Disposições gerais Regulamento de transporte de doentes', *Diário da República*, 241(Série I), pp. 6084–6095.

República Portuguesa (2016) 'Decreto-Lei n.º 19/2016', *Decreto-Lei n.º 19/2016*. Diário da República n.º 74/2016, Série I de 2016-04-15, n.º 74/201(Série I), pp. 1279–1284.

República Portuguesa (2019) 'Aviso nº 11176/2019', *Diário da República*. iário da República n.º 128/2019, Série II de 2019-07-08, 128(Série II), pp. 19192–19194.

Reuter-Oppermann, M., Van Den Berg, P. L. and Vile, J. L. (2017) 'Logistics for Emergency Medical Service systems', *Health Systems*, 6(3), pp. 187–208. doi: 10.1057/s41306-017-0023-x.

ReVelle, C. *et al.* (1977) 'Facility location: a review of context-free and EMS models.', *Health services research*, 12(2), pp. 129–146. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1071976&tool=pmcentrez&rendertype=abstract.

ReVelle, C. and Marianov, V. (1991) 'A probabilistic FLEET model with individual vehicle reliability requirements', *European Journal of Operational Research*. doi: 10.1016/0377-2217(91)90095-D.

ReVelle, C. S., Eiselt, H. A. and Daskin, M. S. (2008) 'A bibliography for some fundamental problem categories in discrete location science', *European Journal of Operational Research*, 184(3), pp. 817–848. doi: 10.1016/j.ejor.2006.12.044.

ReVelle, C. S. and Hogan, K. (1988) 'A reliability-constrained siting model with local estimates of busy fractions', *Environment and Planning B: Planning and Design*. doi: 10.1068/b150143.

ReVelle, C. S. and Hogan, K. (1989) 'THE MAXIMUM RELIABILITY LOCATION PROBLEM AND α-RELIABLE p -CENTER PROBLEM: DERIVATIVES OF THE PROBABILISTIC LOCATION SET COVERING PROBLEM', *Annals of Operations Research*, 18(1–4), pp. 155–173.

Revelle, C., Schweitzer, J. and Snyder, S. (1996) 'The Maximal Conditional Covering Problem', *INFOR: Information Systems and Operational Research*, 34, pp. 77–91. doi: 10.1080/03155986.1996.11732294.

Revelle, C. and Snyder, S. (1995) 'Integrated fire and ambulance siting: A deterministic model', *Socio-Economic Planning Sciences*. doi: 10.1016/0038-0121(95)00014-3.

ReVelle, C., Toregas, C. and Falkson, L. (1976) 'Applications of the Location Set-covering Problem', *Geographical Analysis*, 8(1), pp. 65–76. doi: 10.1111/j.1538-4632.1976.tb00529.x.

Rosa, J. (2017) *Optimizing Staff Scheduling in Emergency Medical Services: a case at INEM*.

Saydam, C. *et al.* (2013) 'The dynamic redeployment coverage location model', *Health Systems*. doi: 10.1057/hs.2012.27.

Saydam, C. and Aytuğ, H. (2003) 'Accurate estimation of expected coverage: Revisited', *Socio-Economic Planning Sciences*. doi: 10.1016/S0038-0121(02)00004-6.

Schilling, David A *et al.* (1979) 'The Team/Fleet Models for Simultaneous Facility and Equipment Siting', *Transportation Science*, 13(2), pp. 163–175. doi: 10.1287/trsc.13.2.163.

Schilling, David A. *et al.* (1979) 'The Team/Fleet Models for Simultaneous Facility and Equipment Siting', *Transportation Science*. doi: 10.1287/trsc.13.2.163.

Schilling, D. A. *et al.* (1980) 'Some models for fire protection locational decisions', *European Journal of Operational Research*. doi: 10.1016/0377-2217(80)90067-3.

Schmid, V. and Doerner, K. F. (2010) 'Ambulance location and relocation problems with time-dependent travel times', *European Journal of Operational Research*. doi: 10.1016/j.ejor.2010.06.033.

Serra, D. (1996) 'The coherent covering location problem', *Papers in Regional Science: The Journal of the RSAI*. doi: 10.1016/S0966-8349(98)80049-1.

Setzler, H., Saydam, C. and Park, S. (2009) 'EMS call volume predictions: A comparative study', *Computers and Operations Research*. doi: 10.1016/j.cor.2008.05.010.

Shariat-Mohaymany, A. *et al.* (2012) 'Linear upper-bound unavailability set covering models for locating ambulances: Application to Tehran rural roads', *European Journal of Operational Research*. doi: 10.1016/j.ejor.2012.03.015.

Shiah, D. M. and Chen, S. W. (2007) 'Ambulance allocation capacity model', in *HEALTHCOM 2007: Ubiquitous Health in Aging Societies - 2007 9th International Conference on e-Health Networking, Application and Services*. doi: 10.1109/HEALTH.2007.381600.

Silva, F. and Serra, D. (2008) 'Locating emergency services with different priorities: The priority queuing covering location problem', *Journal of the Operational Research Society*. doi: 10.1057/palgrave.jors.2602473.

Smith, H. K., Harper, P. R. and Potts, C. N. (2013) 'Bicriteria efficiency/equity hierarchical location models for public service application', *Journal of the Operational Research Society*. doi: 10.1057/jors.2012.68.

Sorensen, P. and Church, R. (2010) 'Integrating expected coverage and local reliability for emergency medical services location problems', *Socio-Economic Planning Sciences*. doi: 10.1016/j.seps.2009.04.002.

Storbeck, J. E. (1982) 'Slack, natural slack, and location covering', *Socio-Economic Planning Sciences*. doi: 10.1016/0038-0121(82)90020-9.

Su, Q., Luo, Q. and Huang, S. H. (2015) 'Cost-effective analyses for emergency medical services deployment: A case study in Shanghai', *International Journal of Production Economics*. doi: 10.1016/j.ijpe.2015.02.015.

Sung, I. and Lee, T. (2018) 'Scenario-based approach for the ambulance location problem with stochastic call arrivals under a dispatching policy', *Flexible Services and Manufacturing Journal*. doi: 10.1007/s10696-016-9271-5.

Toregas, C. *et al.* (1971) 'The Location of Emergency Service Facilities', *Operations Research*. doi: 10.1287/opre.19.6.1363.

Toro-Díaz, H. *et al.* (2013) 'Joint location and dispatching decisions for Emergency Medical Services', *Computers and Industrial Engineering*. doi: 10.1016/j.cie.2013.01.002.

Toro-Díaz, H. *et al.* (2015) 'Reducing disparities in large-scale emergency medical service systems', *Journal of the Operational Research Society*. doi: 10.1057/jors.2014.83.

Torres, N., Trujillo, L. and Maldonado, Y. (2018) 'Modeling Uncertainty for the Double Standard Model Using a Fuzzy Inference System', *Frontiers in Robotics and AI*, 5, p. 31. doi: 10.3389/frobt.2018.00031.

Totten, V. and Bellou, A. (2013) 'Development of Emergency Medicine in Europe', *Academic Emergency Medicine*, 20, pp. 514–521.

Tribunal de Contas (2010) *Auditoria de Resultados ao Instituto Nacional de Emergência Médica*. Available at: https://www.tcontas.pt/pt/actos/rel_auditoria/2010/2s/audit-dgtc-rel047-2010-2s.pdf.

Tzeng, G. H. and Chen, Y. W. (1999) 'The optimal location of airport fire stations: A fuzzy multi-objective programming and revised genetic algorithm approach', *Transportation Planning and Technology*. doi: 10.1080/03081069908717638.

Vile, J. L. *et al.* (2012) 'Predicting ambulance demand using singular spectrum analysis', *Journal of the Operational Research Society*. doi: 10.1057/jors.2011.160.

Wen, M. and Iwamura, K. (2008) 'Fuzzy facility location-allocation problem under the Hurwicz criterion', *European Journal of Operational Research*. doi: 10.1016/j.ejor.2006.11.029.

Wickham, H. *et al.* (2019) 'Package "dplyr": A Grammar of Data Manipulation'.

Williams, H. P. (2013) *Model Building in Mathematical Programming*. 5th Editio. John Wiley & Sons Ltd. doi: 10.2307/3009304.

Von Winterfeldt, D. and Edwards, W. . (1993) 'Decision analysis and behavioral research', *Advances in Decision Analysis - From Foundations to Applications*. doi: 10.1017/CBO9780511611308.006.

World Health Organization (2005) 'Prehospital Trauma Care Sytems', in *Prehospital Trauma Care Sytems*. doi: 10.1007/s00113-010-1919-0.

Wulterkens, D. (2005) 'EMS in The Netherlands: A Dutch Treat?', *Journal of Emergency Medical Services*.

Yang, L., Jones, B. F. and Yang, S. H. (2007) 'A fuzzy multi-objective programming for optimization of fire station locations through genetic algorithms', *European Journal of Operational Research*. doi: 10.1016/j.ejor.2006.07.003.

Yin, P. and Mu, L. (2012) 'Modular capacitated maximal covering location problem for the optimal siting of emergency vehicles', *Applied Geography*. doi: 10.1016/j.apgeog.2011.11.013.

Zaffar, M. A. *et al.* (2016) 'Coverage, survivability or response time: A comparative study of performance statistics used in ambulance location models via simulation–optimization', *Operations Research for Health Care*. doi: 10.1016/j.orhc.2016.08.001.

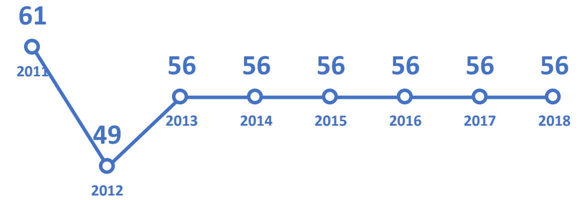
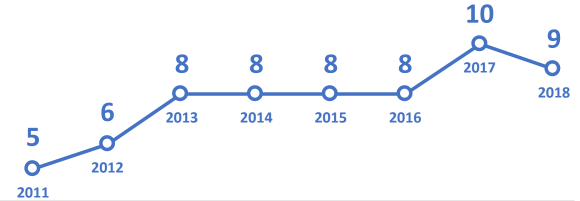
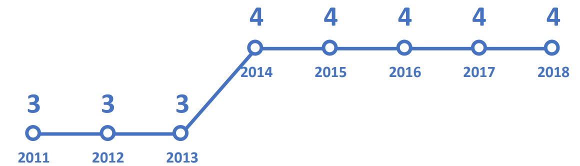
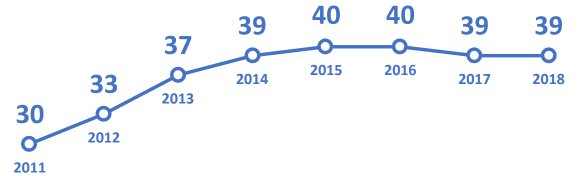
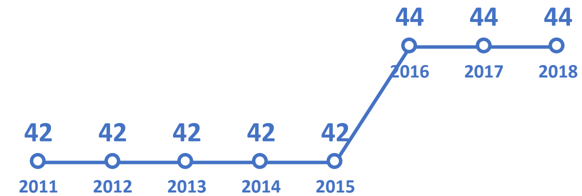
Zhang, Z. H. and Jiang, H. (2014) 'A robust counterpart approach to the bi-objective emergency medical service design problem', *Applied Mathematical Modelling*. doi: 10.1016/j.apm.2013.07.028.

Zhang, Z. H. and Li, K. (2015) 'A novel probabilistic formulation for locating and sizing emergency medical service stations', *Annals of Operations Research*. doi: 10.1007/s10479-014-1758-4.

Zhen, L. *et al.* (2014) 'A simulation optimization framework for ambulance deployment and relocation problems', *Computers and Industrial Engineering*. doi: 10.1016/j.cie.2014.03.008.

APPENDIX A. DESCRIPTION OF SIEM'S EMERGENCY VEHICLES

Table 23 - Emergency vehicles description.

	Name	Description	Available vehicles (2011 – 2018)
	Medical Emergency Ambulance (AEM)	Their function is to promptly transport TEPHs to the emergency scene, stabilize urgent victims and provide transport to the appropriate health facility. They are capable of providing BLS. They are owned and operated by INEM, and an AEM's crew is composed by two TEPHs.	
	Medical Emergency Motorcycle (MEM)	Their function is to promptly transport one TEPH to the emergency scene, stabilize the victim and prepare transportation. MEMs are quicker through traffic, and sometimes perform primary triage as well as assist other vehicles. They are owned and operated by INEM.	
	Inter-Hospital Pediatric Transport Ambulances (TIP)	Their function is to provide stabilization and secondary (inter-hospital) assisted transportation of critical patients aged 0-18 to a differentiated health facility. Their crew is composed of one physician, one nurse and one TEPH and they have the necessary means to stabilize the patient.	
	Immediate Life Support Vehicle (SIV)	They provide differentiated medical care, including resuscitation, defibrillation and medication. They are used to stabilize and transport emergent patients and to perform inter-hospital transportation of critical patients. They are capable of providing Immediate Life Support (ILS). Their crew is composed of a nurse and a TEPH, and they are owned and operated by INEM.	
	Vehicle of Medical Emergency and Reanimation (VMER)	Their function is to transport a medical team directly to an emergency, to stabilize the victim and assist the transportation (although they do not perform the transportation itself). They are capable of providing ALS and their crew includes a physician and a nurse trained in medical emergency services provision working at the emergency department of an NHS hospital. They are owned by INEM.	

Medical Emergency Stations (PEM)

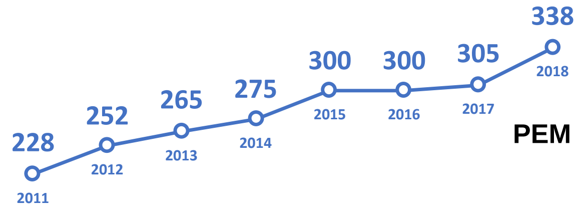
Reserve Station (RES) Ambulances

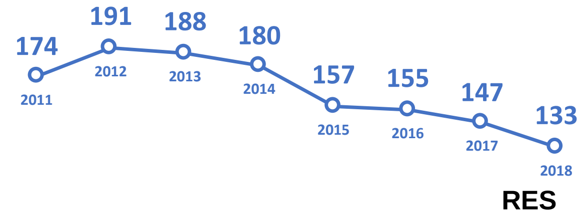
PEM and RES ambulances transport trained professionals in the shortest possible time, to stabilize the victim and provide transportation to the appropriate health facility. They provide BLS and complement other emergency vehicles. They are maintained and operated by a SIEM partner (i.e. firefighters, Red Cross). They follow INEM's protocols and respond when called by a CODU. INEM pays an "exit prize" when an ambulance answers a call (larger for RES).

They are staffed by two members of this SIEM partner, trained in emergency medical care and emergency driving. PEM and RES are established through different collaboration protocols:

- PEM: owned and equipped by INEM and assigned to a SIEM partner. In 2017, INEM has reformulated these agreements so that ambulances are acquired directly by the partner while INEM supports insurance and maintenance costs in installments.
- RES: Completely owned by the SIEM partner, INEM pays an additional "rental" fee.

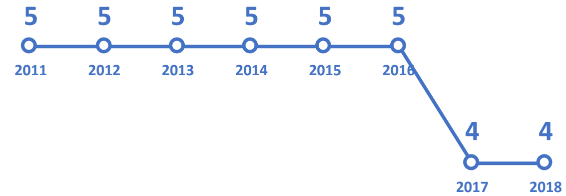
More than one PEM and RES can be located on the same SIEM partner.





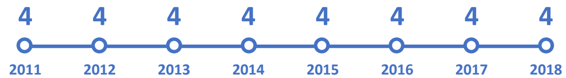
Medical Emergency Helicopter Service (SHEM)

They are used to transport critical patients between health units or between an emergency scene and a health unit, to transport a medical crew to an emergency scene or to transport donated organs. They are equipped with ALS equipment. Their crew includes, besides pilots, a physician and a nurse with specialized training.



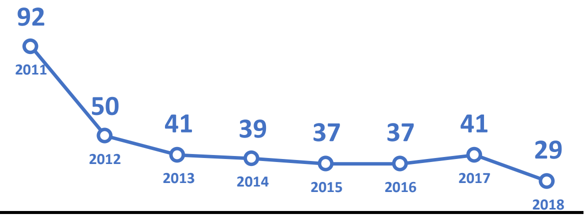
Mobile Unit of Phycological Emergency Intervention (UMIPE)

They intervene in potentially traumatic emergency situations, when the victims' emotional stress requires negotiation to accept help or when children are involved. They are staffed by a TEPH and a psychologist. They are owned and operated by INEM.



Non-INEM Ambulances (NINEM)

These ambulances have the same function as PEM or RES ambulances. However, they are owned and operated by a partner which has no protocol established with INEM. They are used when INEM's ambulances are unavailable or inexistent in the area, or when they are closed to the emergency. Here, INEM pays a single prize every time the vehicle is required.



APPENDIX B. LITERATURE REVIEW SUMMARY TABLE

Table 24 - Literature review summary.

<table>
<tr><th>Section</th><th colspan="2">Approaches</th><th>Reviewed Papers</th></tr>
<tr><td rowspan="2">Early Static Covering Models</td><td colspan="2">Determine number of facilities to provide coverage to all nodes.</td><td>(Toregas et al., 1971; Church and Meadows, 1979)</td></tr>
<tr><td colspan="2">Maximize coverage provided by limited facilities.</td><td>(Church and ReVelle, 1974; Eaton et al., 1985)</td></tr>
<tr><td rowspan="7">Vehicle Unavailability</td><td colspan="2">Multiple coverage models ensuring that more than one vehicle is capable of covering a demand node</td><td>(Daskin and Stern, 1981; Storbeck, 1982; Hogan and ReVelle, 1986; Batta and Mannur, 1990; Church and Gerrard, 2003; Doerner et al., 2005; Degel et al., 2015; Su, Luo and Huang, 2015; Liu et al., 2016)</td></tr>
<tr><td rowspan="5">Probabilistic models: consider explicitly the unavailability of emergency vehicles through the busy fraction (i.e., the fraction of time that it is unavailable to serve demand)</td><td>Reliability Models</td><td>(Chapman and White, 1974; ReVelle and Hogan, 1988, 1989; Marianov and ReVelle, 1994, 1996)</td></tr>
<tr><td>Expected Coverage Models</td><td>(Daskin, 1983; Bianchi and Church, 1988; Repede and Bernardo, 1994; Jayaraman and Srivastava, 1995; Chuang and Lin, 2007)</td></tr>
<tr><td>Hybrid Models</td><td>(Alsalloum and Rand, 2006; Sorensen and Church, 2010)</td></tr>
<tr><td>Queueing-based Models</td><td>(Larson, 1974, 1975; Jarvis, 1985; Bianchi and Church, 1988; ReVelle and Marianov, 1991; Amiri, 1998, 2001; Marianov and Serra, 1998; Saydam and Aytuğ, 2003; Galvão, Chiyoshi and Morabito, 2005; Rajagopalan, Saydam and Xiao, 2008; Ingolfsson, Budge and Erkut, 2008; McLay, 2009)</td></tr>
<tr><td>Site-specific Models</td><td>(ReVelle and Hogan, 1988; Goldberg and Paz, 1991; ReVelle and Marianov, 1991; Ingolfsson, Budge and Erkut, 2008; Knight, Harper and Smith, 2012; Shariat-Mohaymany et al., 2012; Toro-Díaz et al., 2013; Cho et al., 2014; Leknes et al., 2017)</td></tr>
<tr><td colspan="2">Capacitated models: limiting the demand that each vehicle can cover, i.e. establishing capacity constraints</td><td>(Current and Storbeck, 1988; Pirkul and Schilling, 1991; Shiah and Chen, 2007; Schmid and Doerner, 2010; Yin and Mu, 2012)</td></tr>
<tr><td rowspan="9">Demand and Travel Time Uncertainty</td><td rowspan="4">Demand: account for the spatial and temporal uncertainty of emergency requests patterns.</td><td>Queueing Theory (Poisson Process)</td><td>(Batta, June M. Dolan and Krishnamurthy, 1989; Marianov and ReVelle, 1994, 1996; Marianov and Serra, 1998; Borras and Pastor, 2002; Saydam and Aytuğ, 2003; Galvão, Chiyoshi and Morabito, 2005; Cho et al., 2014; Leknes et al., 2017)</td></tr>
<tr><td>Random Variable</td><td>(Beraldi, Bruni and Conforti, 2004; Zhang and Li, 2015; Chu et al., 2018)</td></tr>
<tr><td>Scenarios</td><td>(Beraldi and Bruni, 2009; Berman, Hajizadeh and Krass, 2013; Zhang and Jiang, 2014; Nickel, Reuter-Oppermann and Saldanha-da-Gama, 2016; Sung and Lee, 2018)</td></tr>
<tr><td>Fuzzy Programming</td><td>(Wen and Iwamura, 2008; Torres, Trujillo and Maldonado, 2018)</td></tr>
<tr><td rowspan="5">Travel Time: consider the variability in travel time throughout the planning horizon due to traffic conditions.</td><td>Scenarios</td><td>(Berman, Hajizadeh and Krass, 2013)</td></tr>
<tr><td>Random Variable</td><td>(Aly and White, 1978; Marianov and ReVelle, 1996)</td></tr>
<tr><td>Uncertainty Sets</td><td>(Bertsimas and Ng, 2019)</td></tr>
<tr><td>Coverage Probabilities</td><td>(Daskin, 1987; Goldberg and Paz, 1991; Ingolfsson, Budge and Erkut, 2008; Drezner, Marianov and Wesolowsky, 2016)</td></tr>
<tr><td>Fuzzy Programming</td><td>(Davari, Fazel Zarandi and Hemmati, 2011; Lahijanian, Zarandi and Farahani, 2017) (Torres, Trujillo and Maldonado, 2018)</td></tr>
<tr><td>Time Dependent Models</td><td colspan="2">Account for the dynamic nature of multiple model parameters, including travel time, demand, fleet size and station capacity.</td><td>(Repede and Bernardo, 1994; Rajagopalan, Saydam and Xiao, 2008; Setzler, Saydam and Park, 2009; Cheu, Lei and Aldouri, 2010; Schmid and Doerner, 2010; Başar, Çatay and Ünlüyurt, 2011; Degel et al., 2015; Van Den Berg and Aardal, 2015; Dibene et al., 2017)</td></tr>
<tr><td rowspan="2">Multiple Vehicles and Call Priorities</td><td colspan="2">Multiple Vehicles</td><td>(David A Schilling et al., 1979; David A. Schilling et al., 1979; Charnes and Storbeck, 1980; Revelle and Snyder, 1995; Jayaraman and Srivastava, 1995; Serra, 1996; Amiri, 1998; McLay, 2009; Coskun and Erol, 2010; Davoudpour, Mortaz and Hosseinijou, 2014; Chong, Henderson and Lewis, 2016; Colombo, Cordone and Lulli, 2016; Liu et al., 2016; Van Den Berg, Legemaate and Van Der Mei, 2017)</td></tr>
<tr><td colspan="2">Multiple Call Priorities: consider multiple call types with varying care level needs and coverage thresholds.</td><td>(David A Schilling et al., 1979; Charnes and Storbeck, 1980; Revelle and Snyder, 1995; Silva and Serra, 2008; McLay, 2009; Chong, Henderson and Lewis, 2016; Colombo, Cordone and Lulli, 2016; Liu et al., 2016)</td></tr>
</table>

Alternative Performance Measures	**Alternative Coverage:** try to overcome the limitations of traditional coverage metrics.	**Gradual Coverage**	(Church and Roberts, 1983; Berman and Krass, 2002; Drezner, Wesolowsky and Drezner, 2004; Drezner, Drezner and Goldstein, 2010; Berman and Wang, 2011; Berman, Krass and Wang, 2011; Van Den Berg, Kommer and Zuzáková, 2016)
		Variable Radius	(Berman *et al.*, 2009; Davaria *et al.*, 2010)
		Cooperation	(Berman, Drezner and Krass, 2010a)
	Patient Survival: explicitly consider the medical outcomes of patients as an objective.		(Felder and Brinkmann, 2002; Erkut, Erdogan and Ingolfsson, 2008; McLay and Mayorga, 2010; Knight, Harper and Smith, 2012; Zaffar *et al.*, 2016; Leknes *et al.*, 2017)
	Equity: seek to ensure that the disparities within the system are reduced.		(Drezner, Drezner and Guyse, 2009; McLay and Mayorga, 2010; Chanta *et al.*, 2011; Smith, Harper and Potts, 2013; Chanta, Mayorga and McLay, 2014a, 2014b; Cardoso *et al.*, 2015; Toro-Díaz *et al.*, 2015)
	Multi-Objective Approaches: combine multiple stakeholder perspectives into the optimization models.		(Schilling *et al.*, 1980; Daskin and Stern, 1981; Storbeck, 1982; Revelle, Schweitzer and Snyder, 1996; Tzeng and Chen, 1999; Harewood, 2002; Alsalloum and Rand, 2006; Yang, Jones and Yang, 2007; Kanoun, Chabchoub and Aouni, 2010; Smith, Harper and Potts, 2013; Chanta, Mayorga and McLay, 2014a; Zhang and Jiang, 2014)
Joint Strategic, Tactical and Operational Models	Combine multiple levels of decision making (strategic, tactical and operational) into a single mode.		(Goldberg and Paz, 1991; Borras and Pastor, 2002; Ingolfsson, Budge and Erkut, 2008; Budge, Ingolfsson and Erkut, 2009; Toro-Díaz *et al.*, 2013, 2015; Davoudpour, Mortaz and Hosseinijou, 2014; Grannan, Bastian and McLay, 2014; Ansari, McLay and Mayorga, 2015; Chong, Henderson and Lewis, 2016; Leknes *et al.*, 2017; Sung and Lee, 2018)
Other Issues in EMS Vehicle Planning	**Forecasting:** try to predict demand and travel time for the planning horizon.		(Hall, 1971; Kolesar, Walker and Hausner, 1975; Kvalseth and Deems, 1979; Baker and Fitzpatrick, 1986; McConnel and Wilson, 1998; Channouf *et al.*, 2007; Setzler, Saydam and Park, 2009; Budge, Ingolfsson and Zerom, 2010; Micheletti *et al.*, 2010; Vile *et al.*, 2012)
	Demand Aggregation: group demand points to improve the model's tractability.		(Goodchild, 1979; Francis *et al.*, 2009; Aringhieri *et al.*, 2017)
	Simulation: can be used in simulation-optimization techniques or to validate optimization models.		(Current and Storbeck, 1988; Goldberg *et al.*, 1990; Pirkul and Schilling, 1991; Fu, Glover and April, 2005; Shiah and Chen, 2007; Schmid and Doerner, 2010; Yin and Mu, 2012; Zhen *et al.*, 2014; Aringhieri, Carello and Morale, 2016)
Solution Techniques	**Meta-heuristics**	**Tabu Search**	(Hogan and ReVelle, 1986; Chanta *et al.*, 2011; Saydam *et al.*, 2013; Toro-Díaz *et al.*, 2015)
		Ant Colony	(Doerner *et al.*, 2005)
		Variable Neighbourhood Search	(Schmid and Doerner, 2010)
		Genetic Algorithm	(Saydam and Aytuğ, 2003; Iannoni, Morabito and Saydam, 2009, 2011; Liu *et al.*, 2016)
		Simulated Annealing	(Galvão, Chiyoshi and Morabito, 2005)
	Exact Approaches	**Branch-and-Bound**	(Church and ReVelle, 1974; Church and Meadows, 1979)
		Constraint Generation	(Bertsimas and Ng, 2019)
		Benders Decomposition	(Sung and Lee, 2018)
		Lagrangian Relaxation	(Berman, Hajizadeh and Krass, 2013)
		Single-cut approaches	(Toregas *et al.*, 1971)
		Reduction techniques	(Church and Meadows, 1979).
	Heuristics	**Greedy Heuristics**	(Church and ReVelle, 1974; Silva and Serra, 2008; Iannoni, Morabito and Saydam, 2009, 2011; Berman, Hajizadeh and Krass, 2013; Colombo, Cordone and Lulli, 2016)
		Concentration Heuristic	(Colombo, Cordone and Lulli, 2016)
		Substitution Heuristics	(Batta, June M Dolan and Krishnamurthy, 1989)

APPENDIX C. COMPACT MODEL FORMULATION

Table 25 - Compact model notation, parameters and decision-variables.

Notation	Description	Notation	Description
Sets			
$s \in S$	Working shifts	$p \in P$	Emergency priorities
$t \in T$	Periods in the planning horizon; t= 0 is the beginning of the planning horizon	$v \in V$	Vehicle types
$\|T\|$	Number of planning periods	$l \in L$	Care levels
$d \in D$	Demand points	$f \in F$	Emergency station locations
Subsets			
$f \in F^{exi}$	Existing station locations	$f \in F^{sel}$	Selectable station locations
$f \in F^{new}$	Potential new station locations	$v \in V^{sel}$	Selectable vehicles
Indexed Sets			
$v \in V^f$	Vehicles that can be located at station f	$f \in F^v$	Stations where vehicles v may be located
$v \in V^l$	Vehicles capable of providing care level l	$l \in L^p$	Care levels l required by a call of type p
Parameters			
$OpeningCost_f^t$	Cost of opening station f at the beginning of period t	ε	Minimum fraction of coverage that for a vehicle to be considered as actively cooperating to serve the node
$ClosingCost_f^t$:	Cost of closing a station at site f at the beginning of period t	$StationCap_f^t$	Maximum number of vehicles that can be housed at station f during period t
$CapacityCost_f^t$:	Average cost per vehicle of type v of operating station f during period t	$VeicAva_v^t$	Number of vehicles v available during period t
$OperatingCost_v^{ts}$	Average cost of operating a vehicle of type v during shift s on period t	$MinVeic_{fv}$	Minimum number of vehicles of type v that must be located at station f
$AssignmentCost_{dp}^{ts}$	Average cost of providing care level l to a call of priority p from node d with a vehicle of type v from station f during period t	$InitialVeic_{fv}$	Number of vehicles of type v at station f at the beginning of the planning horizon
Dem_{dp}^{ts}	Requests of priority p from demand node d during on shift s of period t	$MaxVeicShift_v^{ts}$	Number of vehicles v that are available to during shift s of planning period t
$ServiceTime_{dpl}^{ts}$	Average service time to provide care level l to a priority p call from demand node d on shift s of period t	$ShiftLength^s$	Number of time units in shift s
$TravelTime_{fvd}^{ts}$	Travel time for a vehicle of type v from station f to node d on shift s of period t	W_{pl}^1, W_t^2, W_s^3	Weight of covering an emergency of priority p with care l ; during period t and working shift s
γ_{vpl}	1, if a vehicle of type v can provide care level l to call of priority p; 0, otherwise	τ_f	Minimum amount of time that a station at f must remain in operation once opened
$\emptyset_{fvdpl}^{ts}$	Probability that a vehicle of type v departing from a station at site f can cover at care level l a call of priority p from demand node d on shift s of planning period t	$MaxStations^t$, $MaxOpen^t$, $MaxClosed^t$	Maximum number of stations that can be operated/opened/closed during period t
θ	Maximum travel time in the system	i^t	Inflation rate on period t
δ_{fd}	1, if station f can be assigned to calls at node d; 0, otherwise	$Days^t$	Number of days in planning period t
Decision Variables			
$y_f^t \in [0;1]$	1, if an emergency station is operating at site f during planning period t; 0, otherwise.	$ATT_{dpl}^{ts} \in \mathbb{R}_0^+$:	average travel time for care level l to calls p from node d during period t and shift s
$x_{fv}^t \in \mathbb{N}_0$	number of vehicles of type v assigned to a station at site f during planning period t	$MinATT_{dp}^{ts} \in \mathbb{R}_0^+$	average travel time for the first responding vehicle to calls of type p from node d during period t and shift s
$sh_{fv}^{ts} \in \mathbb{N}_0$	number of vehicles of type v assigned to a station at site f during planning period t that are active on working shift s	$E_{fv}^{t+}/E_{fv}^{t-} \in \mathbb{N}_0$:	number of vehicles of type v added to/removed from facility f at the beginning of time period t
$a_{dplfv}^{ts} \in [0;1]$	proportion of demand of priority p from node d for care level l allocated to vehicles of type v at station f during period t and shift s	$H_v^{t+}/H_v^{t-} \in \mathbb{N}_0$	total number of vehicles of type v added to/removed from the fleet at the beginning of period t
$closed_f^t \in [0;1]$, $opened_f^t \in [0;1]$	1, if a station at f is closed/opened at the beginning of planning period t; 0, otherwise.	$b_{fv}^t \in [0;1]$	1, if vehicles of type v are added to site f at the beginning of period t; 0, otherwise.
$w_{dplfv}^{ts} \in [0;1]$	1, if vehicles v at station f are assigned to calls of type p from node d at care level l during shift s and period t	$c_v^t \in [0;1]$	1, if vehicles of type v are added the fleet at the beginning of period t; 0, otherwise.
$\overline{ATT} \in \mathbb{R}_0^+$:	maximum average travel time for the first responding vehicle over the entire region and planning periods	$\beta_{dpl}^{ts} \in [0;1]$	1, if care level l has the shortest average response time for calls of type p from node d during planning period t and working shift s

Objective Functions

$$Z_1 = \max \sum_{p\in P}\sum_{l\in L} W_{pl}^1 \times \sum_{t\in T} W_t^2 \times \sum_{s\in S} W_s^3 \times \sum_{d\in D}\sum_{f\in F}\sum_{v\in V^f} a_{dplfv}^{ts} \times Dem_{dp}^{ts} \times \emptyset_{fvdpl}^{ts}$$

$$\begin{aligned} Z_2 = min \sum_{t\in T\setminus\{0\}} \frac{1}{(1+i^t)^t} & \\ \times \left(\sum_{f\in(F^{exi}\cap F^{sel})} Closed_f^t \times ClosingCost_f^t + \sum_{f\in(F^{new}\cap F^{sel})} Opened_f^t \times OpeningCost_f^t \right) & \\ + \sum_{t\in T} \frac{1}{(1+i^t)^t} \times \sum_{f\in F}\sum_{v\in V^f} x_{fv}^t \times CapacityCost_f^t + \sum_{s\in S} sh_{fv}^{ts} \times OperatingCost_v^{ts} & \\ + \sum_{t\in T} \frac{1}{(1+i^t)^t} \times \sum_{d\in D}\sum_{p\in P}\sum_{l\in L^p}\sum_{s\in S} a_{dplfv}^{ts} \times Dem_{dp}^{ts} \times AssignmentCost_{dplfv}^{ts} & \end{aligned}$$

$$Z_3 = min\ \overline{ATT}$$

Constraints

$$sh_{fv}^{ts} \le x_{fv}^t, \qquad \forall f \in F, v \in V^f, t \in T, s \in S$$

$$\sum_{f\in F^v} x_{fv}^t \le VeicAva_v^t, \qquad \forall\ v \in V, t \in T$$

$$x_{fv}^t \ge MinVeic_{fv}, \qquad \forall\ f \in F, v \in V^f, t \in T\setminus\{0\}$$

$$\sum_{v\in V^f} x_{fv}^t \le StationCap_f^t \times y_f^t, \qquad \forall\ v \in V, t \in T$$

$$\sum_{v\in V^f} x_{fv}^t \ge y_f^t, \qquad \forall\ f \in F, t \in T$$

$$\sum_{f\in F} sh_{fv}^{ts} \ge MaxVeicShift_v^{ts}, \qquad \forall\ v \in V, t \in T, s \in S$$

$$x_{fv}^0 = InitialVeic_{fv}, \qquad \forall\ f \in F^{exi}, v \in V^f$$

$$y_f^0 = 1, \qquad \forall\ f \in F^{exi}$$

$$x_{fv}^t \ge x_{fv}^{t-1}, \qquad \forall\ v \in \overline{V^{sel}}, f \in F^v, t \in T\setminus\{0\}$$

$$y_f^t \ge y_f^{t-1}, \qquad \forall\ f \in \overline{F^{sel}}, t \in T\setminus\{0\}$$

$$y_f^t - y_f^{t-1} = opened_f^t - closed_f^t, \qquad \forall\ f \in F, t \in T\setminus\{0\}$$

$$opened_f^t + closed_f^t \le 1, \qquad \forall\ f \in F, t \in T\setminus\{0\}$$

$$x_{fv}^t - x_{fv}^{(t-1)} = E_{fv}^{t+} - E_{fv}^{t-}, \qquad \forall\ f \in F, v \in V^f, t \in T\setminus\{0\}$$

$$\sum_{f\in F^v} \left(x_{fv}^t - x_{fv}^{(t-1)}\right) = H_v^{t+} - H_v^{t-}, \qquad \forall\ v \in V, t \in T$$

$$E_{fv}^{t+} \le b_{fv}^t \times StationCap_f^t, \qquad \forall\ f \in F, v \in V^f, t \in T\setminus\{0\}$$

$$E_{fv}^{t-} \le (1 - b_{fv}^t) \times StationCap_f^{t-1}, \qquad \forall\ f \in F, v \in V^f, t \in T\setminus\{0,1\}$$

$$E_{fv}^{1-} \le (1 - b_{fv}^1) \times InitialVeic_{fv}, \qquad \forall\ f \in F, v \in V^f$$

$$H_v^{t+} \le c_v^t \times VeicAva_v^t, \qquad \forall\ f \in F, v \in V^f, t \in T\setminus\{0\}$$

$$H_v^{t-} \le (1 - c_v^t) \times VeicAva_f^{t-1}, \qquad \forall\ f \in F, v \in V^f, t \in T\setminus\{0,1\}$$

$$H_v^{1-} \le (1 - c_v^1) \times \sum_{f\in F^v} InitialVeic_{fv}, \qquad \forall\ f \in F, v \in V^f$$

$$\sum_{b=t}^{t+\tau_f^t-1} y_f^b \ge opened_f^t \times \tau_f^t, \qquad \forall\ f \in F, t \in T\setminus\{0\}$$

$$\sum_{f\in F} opened_f^t \le MaxOpen^t, \qquad \forall\ t \in T\setminus\{0\}$$

$$\sum_{f\in F} closed_f^t \leq MaxClosed^t\,, \qquad \forall\, t \in T\backslash\{0\}$$

$$\sum_{f\in F} y_f^t \leq MaxStations^t\,, \qquad \forall\, t \in T\backslash\{0\}$$

$$\left(\sum_{f\in F} E_{fv}^{t+}\right) - H_v^{t+} \leq MaxReal_v^t, \qquad \forall\, v \in V, t \in T$$

$$\sum_{f\in F}\sum_{v\in V^f} a_{dplfv}^{ts} \leq 1, \qquad \forall\, d \in D, p \in P, l \in L^p, s \in S, t \in T$$

$$\sum_{l\in L^p} a_{dplfv}^{ts} \leq sh_{fv}^{ts}, \qquad \forall\, d \in D, p \in P, f \in F, v \in V, s \in S, t \in T$$

$$a_{dplfv}^{ts} \leq \gamma_{vpl} \times\ \delta_{fd}, \qquad \forall \in D, p \in P, l \in L, t\, f \in F, v \in V, s \in S, t \in T$$

$$w_{dplfv}^{ts} \geq a_{dplfv}^{ts}\,, \qquad \forall \in D, p \in P, l \in L, t\, f \in F, v \in V, s \in S, t \in T$$

$$\varepsilon \times w_{dplfv}^{ts} \leq a_{dplfv}^{ts}\,, \qquad \forall \in D, p \in P, l \in L, t\, f \in F, v \in V, s \in S, t \in T$$

$$\sum_{d\in D}\sum_{p\in P}\sum_{l\in L} Dem_{dp}^{ts} \times a_{dplfv}^{ts} \times \left(TravelTime_{fvd}^{ts} + ServiceTime_{dpl}^{ts}\right)$$
$$\leq \rho^{max} \times sh_{fv}^{ts} \times ShiftLength^s, \qquad \forall\, f \in F, v \in V^f, s \in S, t \in T$$

$$\sum_{f\in F}\sum_{v\in V^f} w_{dplfv}^{ts}\ \geq N\ , \qquad \forall\, d \in D, p \in P, l \in L^p, s \in S, t \in T$$

$$ATT_{dpl}^{ts} = \left(\sum_{f\in F}\sum_{v\in V^f} a_{dplfv}^{ts}\ \times\ TravelTime_{fvd}^{ts}\right) + \left(1 - \sum_{f\in F}\sum_{v\in V^f} a_{dplfv}^{ts}\right) \times \theta,$$
$$\forall\, d \in D, p \in P, l \in L^p, s \in S, t \in T$$

$$MinATT_{dp}^{ts} \geq\ ATT_{dpl}^{ts}\ - \theta \times \beta_{dpl}^{ts}, \qquad \forall\, d \in D, p \in P, l \in L^p, s \in S, t \in T$$

$$\sum_{l\in L^p} \beta_{dpl}^{ts} =\ |L^p| - 1\,, \qquad \forall\, d \in D, p \in P, l \in L^p, s \in S, t \in T$$

$$\overline{ATT}\ \geq MinATT_{dp}^{ts}\,, \qquad \forall\, d \in D, p \in P, s \in S, t \in T$$

SP1:

$$Z_1\ = \max \sum_{p\in P}\sum_{l\in L} W_{pl}^1 \times \sum_{t\in T} W_t^2 \times \sum_{s\in S} W_s^3 \times \sum_{d\in D}\sum_{f\in F}\sum_{v\in V^f} sh_{fv}^{ts} \times Dem_{dp}^{ts} \times \emptyset_{fvdpl}^{ts}$$

Subject to:

$$sh_{fv}^{ts} \leq x_{fv}^t\,, \qquad \forall f \in F, v \in V^f, t \in T, s \in S$$

$$\sum_{f\in F^v} x_{fv}^t \leq VeicAva_v^t\,, \qquad \forall\, v \in V, t \in T$$

$$x_{fv}^t \geq MinVeic_{fv}\,, \qquad \forall\, f \in F, v \in V^f, t \in T\backslash\{0\}$$

$$\sum_{v\in V^f} x_{fv}^t \leq StationCap_f^t \times y_f^t\,, \qquad \forall\, v \in V, t \in T$$

$$\sum_{v\in V^f} x_{fv}^t \geq y_f^t\,, \qquad \forall\, f \in F, t \in T$$

$$\sum_{f\in F} sh_{fv}^{ts} \geq MaxVeicShift_v^{ts}\,, \qquad \forall\, v \in V, t \in T, s \in S$$

$$x_{fv}^0 = InitialVeic_{fv}\,, \qquad \forall\, f \in F^{exi}, v \in V^f$$

$$y_f^0 = 1\,, \qquad \forall\, f \in F^{exi}$$

$$x_{fv}^t \geq x_{fv}^{t-1}\,, \qquad \forall\, v \in \overline{V^{sel}}, f \in F^v, t \in T\backslash\{0\}$$

$$y_f^t \geq y_f^{t-1}\,, \qquad \forall\, f \in \overline{F^{sel}}, t \in T\backslash\{0\}$$

$$y_f^t - y_f^{t-1} = opened_f^t - closed_f^t\,, \qquad \forall\, f \in F\,, t \in T\backslash\{0\}$$

$$opened_f^t + closed_f^t \le 1\,, \qquad \forall\, f \in F\,, t \in T\backslash\{0\}$$

$$x_{fv}^t - x_{fv}^{(t-1)} = E_{fv}^{t+} - E_{fv}^{t-}\,, \qquad \forall\, f \in F, v \in V^f, t \in T\backslash\{0\}$$

$$\sum_{f \in F^v} \left(x_{fv}^t - x_{fv}^{(t-1)}\right) = H_v^{t+} - H_v^{t-}, \qquad \forall\, v \in V, t \in T$$

$$E_{fv}^{t+} \le b_{fv}^t \times StationCap_f^t\,, \qquad \forall\, f \in F\,, v \in V^f, t \in T\backslash\{0\}$$

$$E_{fv}^{t-} \le (1 - b_{fv}^t) \times StationCap_f^{t-1}\,, \qquad \forall\, f \in F\,, v \in V^f, t \in T\backslash\{0,1\}$$

$$E_{fv}^{1-} \le (1 - b_{fv}^1) \times InitialVeic_{fv}\,, \qquad \forall\, f \in F\,, v \in V^f$$

$$H_v^{t+} \le c_v^t \times VeicAva_v^t\,, \qquad \forall\, f \in F\,, v \in V^f, t \in T\backslash\{0\}$$

$$H_v^{t-} \le (1 - c_v^t) \times VeicAva_f^{t-1}\,, \qquad \forall\, f \in F\,, v \in V^f, t \in T\backslash\{0,1\}$$

$$H_v^{1-} \le (1 - c_v^1) \times \sum_{f \in F^v} InitialVeic_{fv}\,, \qquad \forall\, f \in F\,, v \in V^f$$

$$\sum_{b=t}^{t+\tau_f^t-1} y_f^b \ge opened_f^t \times \tau_f^t\,, \qquad \forall\, f \in F, t \in T\backslash\{0\}$$

$$\sum_{f \in F} opened_f^t \le MaxOpen^t, \qquad \forall\, t \in T\backslash\{0\}$$

$$\sum_{f \in F} closed_f^t \le MaxClosed^t\,, \qquad \forall\, t \in T\backslash\{0\}$$

$$\sum_{f \in F} y_f^t \le MaxStations^t\,, \qquad \forall\, t \in T\backslash\{0\}$$

$$\left(\sum_{f \in F} E_{fv}^{t+}\right) - H_v^{t+} \le MaxReal_v^t, \qquad \forall\, v \in V, t \in T$$

SP2:

$$Z_1 = \max \sum_{p \in P} \sum_{l \in L} W_{pl}^1 \times \sum_{t \in T} W_t^2 \times \sum_{s \in S} W_s^3 \times \sum_{d \in D} \sum_{f \in F} \sum_{v \in V^f} a_{dplfv}^{ts} \times Dem_{dp}^{ts} \times \emptyset_{fvdpl}^{ts}$$

Subject to:

$$\sum_{f \in F} \sum_{v \in V^f} a_{dplfv}^{ts} \le 1, \qquad \forall\, d \in D, p \in P, l \in L^p, s \in S, t \in T$$

$$\sum_{l \in L^p} a_{dplfv}^{ts} \le sh_{fv}^{ts}, \qquad \forall\, d \in D, p \in P, f \in F, v \in V, s \in S, t \in T$$

$$a_{dplfv}^{ts} \le \gamma_{vpl} \times \delta_{fd}, \qquad \forall \in D, p \in P, l \in L, t\, f \in F, v \in V, s \in S, t \in T$$

$$w_{dplfv}^{ts} \ge a_{dplfv}^{ts}\,, \qquad \forall \in D, p \in P, l \in L, t\, f \in F, v \in V, s \in S, t \in T$$

$$\varepsilon \times w_{dplfv}^{ts} \le a_{dplfv}^{ts}\,, \qquad \forall \in D, p \in P, l \in L, t\, f \in F, v \in V, s \in S, t \in T$$

$$\sum_{d \in D} \sum_{p \in P} \sum_{l \in L} Dem_{dp}^{ts} \times a_{dplfv}^{ts} \times \left(TravelTime_{fvd}^{ts} + ServiceTime_{dpl}^{ts}\right)$$
$$\le \rho^{max} \times sh_{fv}^{ts} \times ShiftLength^s, \qquad \forall\, f \in F, v \in V^f, s \in S, t \in T$$

$$\sum_{f \in F} \sum_{v \in V^f} \sum_{l \in L^p} w_{dplfv}^{ts} \ge N\,, \qquad \forall\, d \in D, p \in P, s \in S, t \in T$$

APPENDIX D. EMERGENCY REQUEST ANALYSIS

In order to verify the hypothesis that the volume of emergency requests varies strongly with time of the day and month, a Kruskal-Wallis test is conducted on the average number of calls per day and shift. For that purpose, the mean and standard deviation of the number of calls per day in each month and shift are calculated in R, and are presented in Table 26.

Table 26 - Mean and standard deviation (SD) of the number of calls per day per month and shift.

	Lisbon				Setúbal			
	P1		P3		P1		P3	
Month	**Average**	**SD**	**Average**	**SD**	**Average**	**SD**	**Average**	**SD**
Jan-17	26.16129	3.899545	175.6452	22.75895	4.54839	2.39219	175.64520	22.75895
Feb-17	23.57143	4.220133	160.2143	19.03283	4.28571	1.88281	160.21430	19.03283
Mar-17	22.77419	5.277055	159.9032	13.93881	3.82759	1.64900	159.90320	13.93881
Apr-17	19.16667	4.25954	154.7333	18.75419	3.82759	1.77420	154.73330	18.75419
May-17	19.90323	5.545898	161.7419	15.04431	3.80645	2.19726	161.74190	15.04431
Jun-17	20.06667	4.926587	160.8333	21.17429	3.82759	1.62720	160.83330	21.17429
Jul-17	19.77419	4.425002	152.4194	15.44188	3.27586	1.38607	152.41940	15.44188
Aug-17	17.54839	3.631700	143.7419	15.39257	3.10000	1.56139	143.74190	15.39257
Sep-17	19.13333	4.932184	161.8667	18.3091	3.67857	1.56474	161.86670	18.30910
Oct-17	21.45161	5.091422	165.4194	18.79499	3.93548	2.26474	165.41940	18.79499
Nov-17	23.93333	4.912569	168.8667	19.21876	4.03333	1.93842	168.86670	19.21876
Dec-17	23.87097	4.455274	170.4839	18.10133	4.70968	1.84740	170.48390	18.10133
Jan-18	24.87097	5.327551	171.871	22.05409	4.54839	2.26331	171.87100	22.05409
Feb-18	27.17857	6.359907	180.0000	15.87217	4.46429	2.25228	180.00000	15.87217
Mar-18	24.83871	4.747608	168.7742	17.23313	3.96667	2.20475	168.77420	17.23313
Apr-18	24.03333	4.634826	164.2000	21.11773	3.96429	2.38020	164.20000	21.11773
May-18	20.6129	4.499462	166.7742	18.59339	3.50000	1.81557	166.77420	18.59339
Jun-18	20.56667	4.438727	165.2667	16.88487	3.34483	1.56470	165.26670	16.88487
Shift								
Morning	8.02564	2.96761	58.62454	9.58344	1.89781	1.07716	9.18681	3.06712
Evening	9.86264	3.49645	76.26557	14.15021	2.09469	1.22769	12.01282	3.62886
Night	4.50000	2.13070	29.11883	8.17072	1.44211	0.69276	4.75506	2.38167

By analysing these values, it becomes apparent that the volume of calls is not steady, as confirmed by the test results presented in section 6.6.

When choosing a statistical model to help predict the volume of calls in the planning horizon, besides analysing the inter-call times, the number of requests per day is also analysed. By evaluating the histograms of the number of emergencies per day (as in the Figure bellow), it is concluded that it resembles a Poisson distribution mass function, as would be expected.

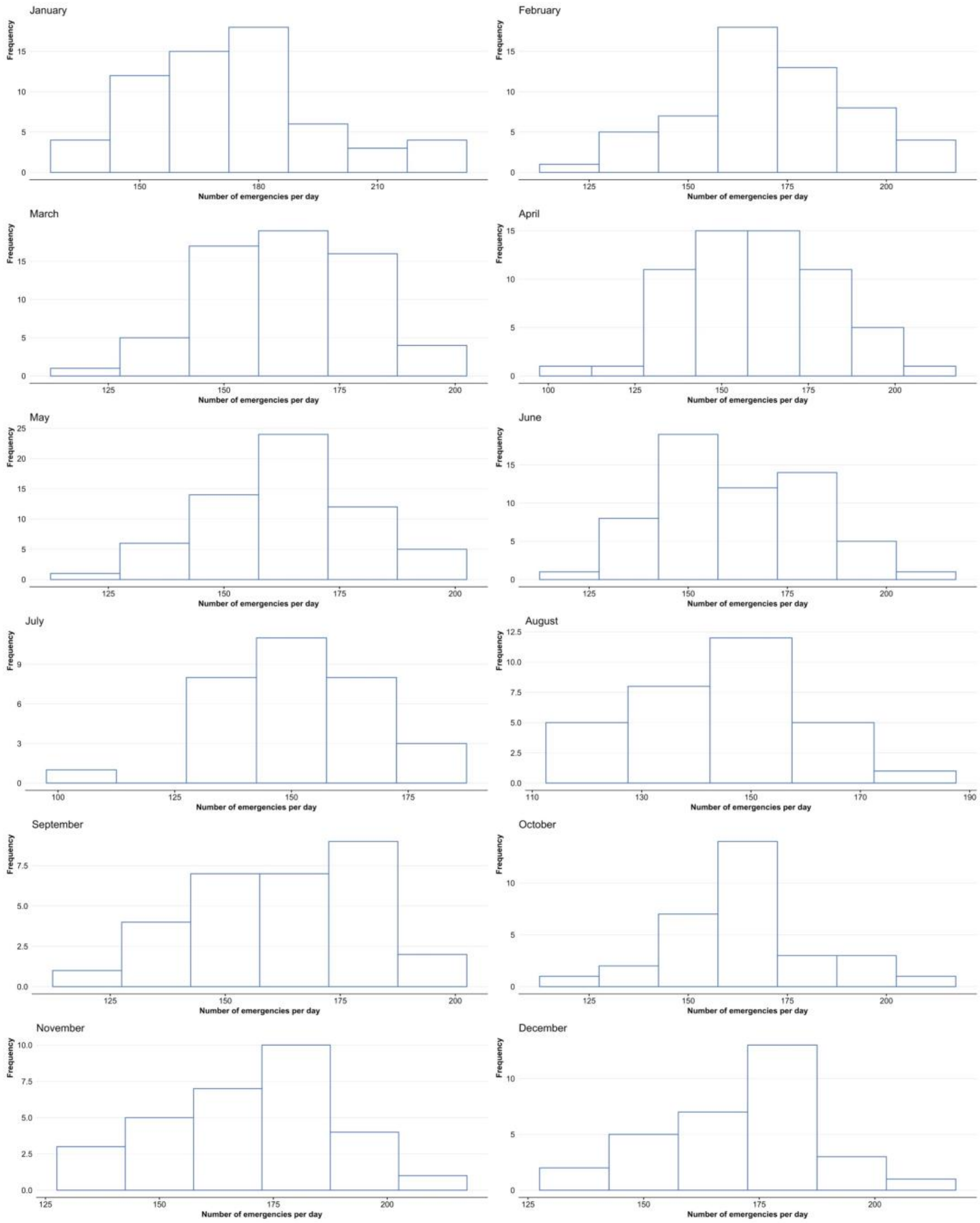


Figure 35 - Histograms of the number of P1 emergencies in Setúbal per day.

APPENDIX E. ADDITIONAL CASE-STUDY DATA

This appendix presents additional case-study data, duly referenced on Chapter 6.

Table 27 - Maximum number of stations that can be opened or closed during the planning period.

	Region	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
OPENING	Lisbon	1	1	1	1	1	2	2	2	2	2	0	1
	Setúbal	1	1	1	1	1	2	2	2	2	2	0	1
CLOSING	Lisbon	1	1	1	1	1	1	1	1	2	1	1	1
	Setúbal	1	1	1	1	1	1	1	1	1	1	1	0

Table 28 - Number of emergency vehicles available on each month for Lisbon and Setúbal.

		Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
LISBON	NINEM	6	6	6	6	6	6	6	6	6	6	6	6
	RES	0	0	0	0	0	0	0	0	0	0	0	0
	PEM	2	2	2	2	2	2	3	3	3	2	2	2
	AEM	14	14	14	14	14	14	14	14	14	14	14	14
	SIV	1	1	1	1	1	1	1	1	1	1	1	1
	VMER	3	3	3	3	3	3	3	3	3	3	3	3
SETÚBAL	NINEM	0	0	0	0	0	0	0	0	0	0	0	0
	RES	0	0	0	0	0	0	1	1	1	0	0	0
	PEM	5	5	5	5	5	5	6	6	6	5	5	5
	AEM	2	2	2	2	2	2	2	2	2	2	2	2
	SIV	0	0	0	0	0	0	0	0	0	0	0	0
	VMER	1	1	1	1	1	1	1	1	1	1	1	1

Table 29 - Availability of emergency vehicles on different shifts in Lisbon and Setúbal.

		January- June | October-December			July-September		
		Morning	Evening	Night	Morning	Evening	Night
LISBON	NINEM	6	6	6	6	6	6
	RES	0	0	0	0	0	0
	PEM	2	2	1	3	3	3
	AEM	14	13	7	14	13	7
	SIV	1	1	1	1	1	1
	VMER	3	3	3	3	3	3
SETÚBAL	NINEM	0	0	0	0	0	0
	RES	0	0	0	0	0	0
	PEM	5	4	3	6	6	6
	AEM	2	2	1	2	2	1
	SIV	0	0	0	0	0	0
	VMER	1	1	1	1	1	1

Table 30 - Fitted arrival rate for P1 and P3 emergencies in Lisbon and Setúbal for each month (emergencies/hour).

			Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
LISBON	**P1**	**Morning**	1.23	1.29	1.29	1.05	1.08	0.99	1.03	1.01	0.98	1.12	1.26	1.20
		Evening	1.21	1.22	1.09	1.03	0.93	0.98	0.99	0.72	0.86	1.00	1.09	1.16
		Night	0.66	0.63	0.55	0.62	0.50	0.55	0.53	0.44	0.53	0.55	0.60	0.58
	P3	**Morning**	9.93	9.65	9.44	9.26	9.51	9.52	8.73	8.29	9.63	9.82	9.53	9.33
		Evening	7.93	8.02	7.60	7.04	7.43	7.05	6.71	6.29	6.99	7.08	8.04	8.01
		Night	3.84	3.57	3.46	3.63	3.57	3.77	3.62	3.39	3.63	3.79	3.51	3.92
SETÚBAL	**P1**	**Morning**	0.20	0.16	0.17	0.17	0.17	0.17	0.15	0.15	0.14	0.14	0.14	0.21
		Evening	0.23	0.24	0.21	0.21	0.25	0.14	0.19	0.17	0.19	0.25	0.24	0.19
		Night	0.16	0.18	0.14	0.13	0.11	0.15	0.15	0.11	0.17	0.15	0.18	0.19
	P3	**Morning**	1.49	1.53	1.42	1.40	1.49	1.46	1.31	1.21	1.22	1.41	1.35	1.46
		Evening	1.30	1.38	1.20	1.28	1.18	1.24	1.19	1.21	1.19	1.25	1.22	1.19
		Night	0.67	0.62	0.63	0.59	0.64	0.70	0.59	0.54	0.62	0.57	0.55	0.60

Table 31 - Maximum travel times for different cluster partitions and shifts.

	Cluster Partition	Shift Morning	Evening	Night
LISBON	1% (33)	53.0508206	48.9448809	39.7457861
	2% (66)	47.4069924	43.8451881	32.3604476
	5% (165)	51.5477667	46.1591022	45.5677732
SETÚBAL	1% (27)	37.5065337	31.5487246	29.4652411
	2% (54)	39.6760742	31.4455207	28.8048597
	5% (135)	39.5771775	32.6457011	28.6140990

Table 32 - Estimated aggregate service times for Setúbal.

	P1 - ALS			P1 - BLS			P3 - BLS		
Month	**Morning**	**Evening**	**Night**	**Morning**	**Evening**	**Night**	**Morning**	**Evening**	**Night**
January	60.89	55.56	58.28	58.97	65.46	63.04	63.85	65.37	61.10
February	58.85	59.97	53.71	71.40	66.79	59.53	62.37	65.46	67.56
March	55.65	43.34	64.65	66.90	62.14	67.80	62.50	62.97	65.72
April	52.64	55.78	58.07	60.33	64.68	74.19	61.98	62.60	66.08
May	61.42	56.75	58.23	65.38	71.42	74.41	61.13	61.71	65.75
June	57.34	43.49	56.08	68.53	67.10	69.11	62.31	61.14	65.17
July	67.13	60.84	53.89	61.96	67.91	68.93	59.45	64.17	62.71
August	57.04	54.79	56.60	67.39	69.79	72.55	57.52	66.16	64.69
September	51.45	53.25	58.09	54.97	72.22	81.33	57.82	63.97	61.10
October	60.32	60.94	56.47	63.09	68.51	74.84	61.32	61.81	68.36
November	55.29	41.45	57.72	63.47	61.55	73.51	60.64	61.93	66.37
December	58.80	57.14	46.32	72.00	70.51	67.28	62.58	60.47	65.57

Table 33 - Example of coverage probabilities for three stations and three demand regions (of the 1% cluster partitioning) in Setúbal.

	Demand Areas			
Station	**0**	**1**	**2**	**3**
BV Setúbal - Azeitão	0.0092	0.8976	0.1382	0.0164
BV Setúbal - Sede	0.1344	0.0188	0.9922	0.7920
Cruz Vermelha Setúbal	0.4410	0.0108	0.9914	0.6080

APPENDIX F. SCALABILITY ANALYSIS RESULTS

In order to assess the scalability of the implemented model, several instances of different sizes are analysed. The number of emergency priorities, time periods, shifts and care levels is fixed, while the number of stations and demand areas are variable. The analysis uses the base line instances, whose size is presented in Table 34. Each test is represented by a letter – L for Lisbon, S for Setúbal – and two numbers: the first represents the size of the instance (0 for the 1% partition, 1 for the 2% partition and 2 for the 5% partition), and the second is the experiment number (see section 7.3).

Table 34 - Size of the base-line test instances.

Instance	City	Demand areas	Stations	Constraints	Variables		
					Binary	Integer	Continuous
L0.0	Lisbon	33	42	866430	212586	4178	213841
L1.0	Lisbon	66	42	1716978	420666	4178	427681
L2.0	Lisbon	165	42	4268862	1044366	4178	1069201
S0.0	Setúbal	27	16	171189	39106	1016	42769
S1.0	Setúbal	54	16	344565	78418	1016	87121
S2.0	Setúbal	135	16	852297	193546	1016	217009

A Lexicographic Ordering approach is used to handle multiple objectives. For this purpose, objectives are ranked according to their priority for INEM. The resulting hierarchy, in decreasing order of importance, is Z1, Z2 and Z3. Three runs are conducted for each instance. Firstly, coverage is maximized (run 0). Subsequently, the value of this objective is bounded by a constraint (lower bound) and total costs are minimized (run 1). In the third run, an upper bound on costs and a lower bound on coverage are set and equity is maximized by minimizing $Z3$ (run 2). Additionally, the solution of the previous run is used as a MIP start solution in CPLEX, given that it is always a feasible solution.

Before proceeding to assess the computational time of these instances, it is interesting to analyze the rate of convergence to optimality of the branch-and-cut procedure for two test instances (Figure 36). It is possible to conclude that although a tight upper-bound is quickly set by CPLEX, and a quality feasible solution is found in some minutes, ensuring optimality is challenging. This suggests that the Linear Relaxation of the proposed model is relatively tight.

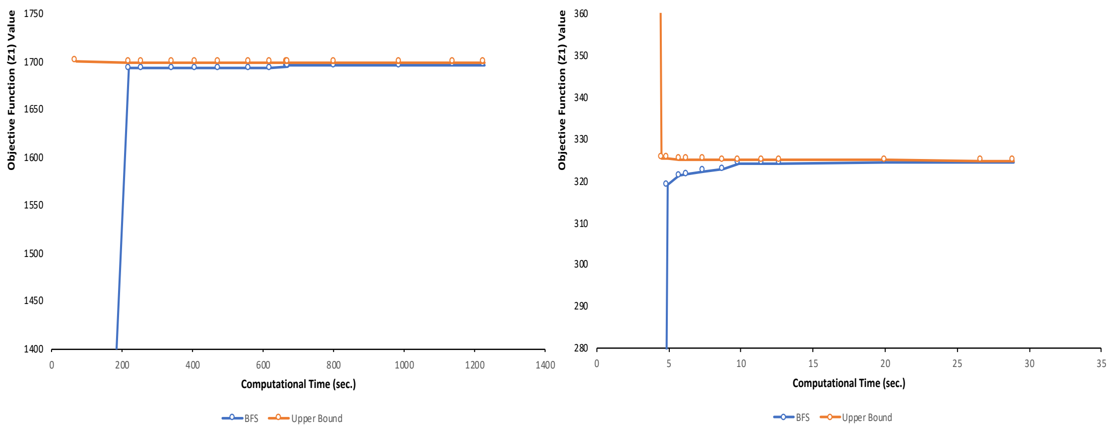


Figure 36 - Convergence of the branch-and-cut procedure in the base-line instances (Left: Lisbon, 1% cluster partition; Right: Setúbal, 5% cluster partition).

Since proving optimality requires a significant amount of computational time without great improvement in the objective function, all tests are conducted using an optimality gap of 0.5% for the Z1 objective (run 0), while in the subsequent runs, a relative gap of 5% is applied. Different relative gaps are used because the first objective is more imperative and, thus, greater care should be taken when optimizing it. Computational times for run 0 and objective function values are presented in Table 35.

Table 35 - Computational time and objective function results for the model scalability analysis.

Instance	Run 0				Run 1		Run 2	
	Coverage (Z1)	**Gap** (%)	**Branch-and-Cut Time (s)**	**Total Time (s)**	**Cost** (Z2)	**Gap** (%)	**Equity** (Z3)	**Gap** (%)
L0.0	1691.431	0.47%	188.7	237.28	65011.790	2.68%	15.105	0.00%
L1.0	1647.422	0.44%	874.19	971.396	65131.897	4.24%	15.471	0.00%
L2.0	1652.116	0.18%	8821.17	9191.09	65934.688	4.93%	15.864	0.00%
S0.0	323.4685	0.00%	5.50	22.44	15157.960	3.20%	16.341	0.00%
S1.0	322.4058	0.00%	18.08	51.19	13038.877	4.17%	17.476	0.00%
S2.0	322.8103	0.00%	179.39	247.64	13986.250	4.60%	16.686	0.00%

As would be expected, the computational effort required to solve the model increases exponentially with the size of the instance. This is mainly due to the fact that, being a dynamic model with micro and macro time periods, the number of variables and constraints increases significantly even for small networks. In fact, by considering 12 time periods and three shifts, the number of variables (which are indexed by these parameters) increases 36 times when compared to a static model.

The computational burden also increases significantly in the second and, mostly, third run of the Lexicographic method. This is because the additional constraints related with the previous objectives render the model much more difficult, even though an initial feasible solution is always provided by the prior run. Without this initial solution, the computational time is even greater. Overall, the computational overhead of the model is significant, since it is a large combinatorial problem. Nevertheless, given that speed is not crucial in this study, CPLEX is used in all experiences except when stated otherwise.

The results also suggest that the chosen cluster partition only slightly influences the results. As mentioned, larger aggregation areas introduce errors in the model. Therefore, the 5% partitions are expected to be more accurate than the 1% partitions. In the tested instances, using larger aggregation areas results in an overestimation of expected coverage of 2.38% in Lisbon and 5.86% in Setúbal, between the 1% and 5% cluster partitions. Contrarily, the equity objective presents an underestimation of 4.78% in Lisbon and 2.07% in Setúbal. Regarding costs, the results are not conclusive. In Lisbon, they are underestimated by 1.40% in Lisbon, while in Setúbal they are overestimated by 8.37%.

In light of these results, and in order to reduce the computational cost and allow for more experiments to be carried out, the remaining experiences are conducted using the L0 instances for Lisbon. For Setúbal, the S2 test instances are used since the computational cost is still acceptable for all runs. A relative gap of 0.5% is used when maximizing expected coverage, while a gap of 5% is applied when optimizing the remaining objectives. Additionally, a time limit of 24 hours is used for each run. The following sections focus on exploring the model to assess the impact of alternative policies in the SIEM's expected performance.

APPENDIX G. SAMPLE MODEL SOLUTION

Table 36 - Proposed solution sample for Lisbon in the base-line scenario.

Station	Vehicle	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
INEM Sede	**AEM**	(1, 1, 1)	(1, 1, 1)	(2, 2, 2)	(2, 2, 1)	(2, 2, 1)	(2, 2, 1)	(2, 1, 2)	(1, 1, 0)	(2, 2, 2)	(1, 1, 1)	(1, 1, 1)	(2, 2, 2)
GNR Reg. Cavalaria - Ajuda	**AEM**	(1, 1, 1)	(1, 1, 0)	(1, 0, 0)	(1, 1, 0)	(1, 1, 0)	(1, 1, 1)	(1, 1, 0)	(1, 1, 0)	(1, 1, 0)	(1, 1, 1)	(1, 1, 0)	(1, 0, 0)
Esquadra PSP - Bº Boavista	**AEM**	(1, 1, 1)	(1, 1, 0)	(1, 1, 0)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 0)	(1, 1, 1)	(1, 1, 0)	(1, 1, 0)	(1, 1, 1)	(1, 1, 0)
Centro Saúde Lóios - Olivais	**AEM**	(3, 3, 0)	(1, 1, 1)	(1, 1, 1)	(2, 2, 1)	(1, 1, 0)	(2, 2, 1)	(1, 1, 0)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)
Hosp. Curry Cabral	**AEM**	(2, 2, 2)	(1, 1, 1)	(1, 1, 0)	(1, 1, 1)	(1, 1, 1)	(1, 1, 0)	(1, 1, 0)	(3, 3, 1)	(3, 2, 0)	(3, 3, 1)	(3, 3, 1)	(1, 1, 0)
Escola S.D. Benfica	**AEM**	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(2, 2, 1)	(1, 1, 1)	(1, 1, 0)	(1, 1, 1)	(1, 1, 0)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 0)
GNR Brigada Fiscal - Beato	**AEM**	(1, 1, 0)	(1, 1, 0)	(1, 1, 0)	(1, 1, 0)	(1, 1, 0)	(1, 1, 0)	(1, 1, 0)	(1, 1, 0)	(1, 1, 0)	(1, 1, 0)	(1, 1, 1)	(1, 1, 1)
GNR BT - Alcântara	**AEM**	(1, 1, 0)	(1, 1, 0)	(1, 1, 0)	(1, 1, 0)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 0)	(1, 1, 0)	(1, 1, 0)	(1, 1, 0)	(1, 1, 1)
Hosp. Egas Moniz	**AEM**	(1, 1, 0)	(1, 0, 0)	(1, 1, 0)	(1, 0, 1)	(1, 1, 0)	(1, 1, 0)	(1, 1, 0)	(1, 1, 0)	(1, 1, 0)	(1, 1, 0)	(0, 0, 0)	(1, 1, 0)
Hosp. Egas Moniz	**SIV**	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(1, 0, 0)	(0, 0, 0)
Hosp. São Francisco Xavier	**VMER**	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)
Hosp. São José	**AEM**	(0, 0, 0)	(1, 1, 0)	(1, 1, 0)	(1, 1, 1)	(1, 1, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(1, 1, 0)	(1, 1, 1)
Hosp. São José	**VMER**	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)
Hosp. Santa Maria	**AEM**	(1, 0, 1)	(1, 1, 1)	(1, 1, 1)	(0, 0, 0)	(1, 1, 1)	(1, 1, 1)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(1, 1, 0)	(1, 1, 1)	(1, 1, 0)
Hosp. Santa Maria	**VMER**	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)
Hosp. Santa Maria	**SIV**	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(1, 1, 0)	(1, 1, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)
INEM - R. Infante D. Pedro	**AEM**	(1, 1, 0)	(1, 1, 1)	(0, 0, 0)	(1, 1, 0)	(1, 0, 0)	(1, 1, 1)	(1, 1, 0)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 0)	(1, 1, 1)
INEM - R. Infante D. Pedro	**SIV**	(1, 1, 0)	(1, 1, 0)	(1, 1, 0)	(1, 1, 0)	(1, 1, 0)	(1, 1, 0)	(0, 0, 0)	(0, 0, 0)	(1, 0, 0)	(1, 1, 0)	(0, 0, 0)	(1, 1, 0)
RSB – Av. D. Carlos I	**PEM**	(2, 2, 1)	(2, 2, 1)	(2, 2, 1)	(2, 2, 1)	(2, 2, 1)	(2, 2, 1)	(2, 2, 2)	(2, 2, 2)	(2, 2, 2)	(2, 2, 1)	(2, 2, 1)	(2, 2, 1)
BV Ajuda	**AEM**	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(1, 0, 0)	(0, 0, 0)
BV Ajuda	**NINEM**	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)
BV Beato	**AEM**	(0, 0, 0)	(1, 0, 0)	(1, 1, 0)	(0, 0, 0)	(1, 0, 0)	(0, 0, 0)	(1, 1, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)
BV Beato	**NINEM**	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)
BV Campo de Ourique	**NINEM**	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)
BV Cabo Ruivo	**AEM**	(0, 0, 0)	(1, 1, 0)	(1, 1, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(1, 0, 1)	(0, 0, 0)	(0, 0, 0)
BV Cabo Ruivo	**NINEM**	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)
BV Lisboa	**AEM**	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(1, 1, 0)
BV Lisboa	**NINEM**	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)
BV Lisbonenses	**NINEM**	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)
Centro de Saúde Graça	**AEM**	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 1, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)
Centro de Saúde Marvila	**AEM**	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(1, 1, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)
Centro de Saúde Alvalade	**AEM**	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(1, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)
Centro de Saúde Benfica	**AEM**	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(1, 1, 1)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)
Hospital Cruz Vermelha	**AEM**	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(1, 1, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)

Note: The solution for each facility/month is represented by (M, E, N), where M is the number of active vehicles during the morning, E is the number of active vehicles during the evening and N is the number of active vehicles during the night.

APPENDIX H. COMPUTATIONAL RESULTS

Table 37 - Computational results summary.

	Instance Characteristics					Instance Size					Computational Time		Objectives					
				Demand		Variables				Non-zero		Total Time					Equity	
Experiment	Instance	Run	City	areas	Stations	Binary	Integer	Continuous	Constraints	coefficients	Root + BC (s)	(s)	Coverage (Z1)	Gap (%)	Cost (Z2)	Gap (%)	(Z3)	Gap (%)
Toy Instance	T0	0	Toy Instance	3	3	177	51	235	1035	2464	0.09	3.68	5457.053	0.00%	—	—	—	—
	T0	1	Toy Instance	3	3	177	51	235	1034	2344	0.11	4.91	5457.053	0.00%	4946.559	0.01%	—	—
	T0	2	Toy Instance	3	3	177	51	235	1035	2464	0.16	4.95	5457.053	0.00%	4946.559	0.01%	0.496	0.00%
	T1	0	Toy Instance	6	6	2696	412	2971	12604	38648	0.50	5.25	8796.711	0,00%	—	—	—	—
	T1	1	Toy Instance	6	6	2696	412	2971	12605	40838	0.33	2.21	8796.711	0,00%	37089.903	0.94%	—	—
	T1	2	Toy Instance	6	6	2696	412	2971	12606	41158	1.80	3.49	8796.711	0,00%	37089.903	0.94%	0.548	4.36%
0. Base Line Optimization & Scalability Analysis	L0.0	0	Lisbon	33	42	212586	4178	213841	866430	3129900	188.7	237.28	1691.809	0.45%	—	—	—	—
	L0.0	1	Lisbon	33	42	212586	4178	213841	866431	3129900	130.69	134.062	1691.809	0.45%	65011.790	2.68%	—	—
	L0.0	2	Lisbon	33	42	212586	4178	213841	866432	3284750	6956.99	7039.87	1691.809	0.45%	65011.790	2.68%	15.105	0.00%
	L1.0	0	Lisbon	66	42	420666	4178	427681	1716978	5726482	2207.77	2300.332	1647.423	0.44%	—	—	—	—
	L1.0	1	Lisbon	66	42	420666	4178	427681	1716979	6079522	874.19	971.396	1647.423	0.44%	65131.897	4.24%	—	—
	L1.0	2	Lisbon	66	42	420666	4178	427681	1716980	6360219	12072.97	12166.612	1647.423	0.44%	65131.897	4.24%	15.471	0.00%
	L2.0	0	Lisbon	165	42	1044366	4178	1069201	4268862	13700185	9137.39	9310.026	1651.999	0.18%	—	—	—	—
	L2.0	1	Lisbon	165	42	1044366	4178	1069201	4268863	14401474	6282.66	7335.943	1651.999	0.18%	65934.688	4.93%	—	—
	L2.0	2	Lisbon	165	42	1044366	4178	1069201	4268864	14958264	13564.98	13865.731	1651.999	0.18%	65934.688	4.93%	15.864	0.00%
	S0.0	0	Setúbal	27	10	39106	1016	42769	171189	483332	5.5	22.44	323.394	0.49%	—	—	—	—
	S0.0	1	Setúbal	27	10	39106	1016	42769	171190	500424	124.77	141.31	323.394	0.49%	15157.960	3.20%	—	—
	S0.0	2	Setúbal	27	10	39106	1016	42769	171191	513698	533.36	580.21	323.394	0.49%	15157.960	3.20%	16.341	0.00%
	S1.1	0	Setúbal	54	10	78418	1016	87121	344565	948088	18.08	51.19	322.406	0.00%	—	—	—	—
	S1.1	1	Setúbal	54	10	78418	1016	87121	344566	974796	232.86	267.08	322.406	0.00%	13038.877	4.17%	—	—
	S1.1	2	Setúbal	54	10	78418	1016	87121	344567	995408	236.48	141.8	322.406	0.00%	13038.877	4.17%	17.476	0.00%
	S2.1	0	Setúbal	135	10	193546	1016	217009	852309	2265540	179.39	247.64	321.584	0.00%	—	—	—	—
	S2.1	1	Setúbal	135	10	193546	1016	217009	852310	2306592	1032.45	1104.76	321.584	0.00%	14946.895	4.60%	—	—
	S2.1	2	Setúbal	135	10	193546	1016	217009	852311	2338176	1410.92	1612.09	321.584	0.00%	14946.895	4.60%	17.796	0.00%
1. Current System	L0.1	0	Lisbon	33	20	211362	2486	213841	866430	2722858	15.28	107.67	1686.135	0.10%	—	—	—	—
	L0.1	1	Lisbon	33	20	211362	2486	213841	866431	2916276	79.69	181.19	1686.135	0.10%	66494.661	0.01%	—	—
	L0.1	2	Lisbon	33	20	211362	2486	213841	866432	3068570	514.67	610.61	1686.135	0.10%	66494.661	0.01%	17.856	4.86%
	S2.1	0	Setúbal	135	4	193222	620	217009	852297	2085804	151.58	239.81	320.4299	0.00%	—	—	—	—
	S2.1	1	Setúbal	135	4	193222	620	217009	852298	2126856	144.58	252.6	320.4299	0.00%	17119.534	0.00%	—	—
	S2.1	2	Setúbal	135	4	193222	620	217009	852299	2157828	140.42	219.38	320.4299	0.00%	17119.534	0.63%	18.321	0.48%
2. Seasonal PEMs	L0.2	0	Lisbon	33	42	249018	4877	249481	1011040	3436049	354.73	397.805	1703.41471	0.19%	—	—	—	—
	L0.2	1	Lisbon	33	42	249018	4877	249481	1011041	3663337	214.18	258.981	1703.41471	0.19%	65620.028	3.17%	—	—
	L0.2	2	Lisbon	33	42	249018	4877	249481	1011042	3852777	4624.53	4696.165	1703.41471	0.19%	65620.028	3.17%	15.1046	0.00%
	S2.2	0	Setúbal	135	10	208114	1280	231805	911601	2442939	102.22	174.56	321.583	0.00%	—	—	—	—
	S2.2	1	Setúbal	135	10	208114	1280	231805	911602	2488842	69.14	142.01	321.583	0.00%	14337.553	1.18%	—	—
	S2.2	2	Setúbal	135	10	208114	1280	231805	911603	2527518	366.78	439	321.583	0.00%	14337.553	1.18%	18.233	0.00%
3./4. Fleet Expansion	Results are presented in Table 38																	
5. Impact of Legislation	L0.5	0	Lisbon	33	42	212586	4358	213841	865398	2935450	2597.7	2649.907	1694.51984	0.44%	—	—	—	—
	S2.5	0	Setúbal	135	10	193546	1016	217009	852069	2265300	188.05	259.66	321.584	0.00%	—	—	—	—
6. Vehicle Relocation	L0.6	0	Lisbon	33	42	242444	4360	247105	999340	3402464	10783.8	10818.577	1709.56687	0.35%	—	—	—	—
	S2.6	0	Setúbal	135	10	187745	836	212077	830949	2200680	375.64	443.6	321.494	0.00%	—	—	—	—
7. Green-Field Scenario	L0.7	0	Lisbon	33	42	212586	4358	213841	866088	2935972	82156.05	81751.51	1705.624	0.50%	—	—	—	—
	S2.7	0	Setúbal	135	10	193546	1016	217009	852240	2265439	5533.47	5862.53	322.283	0.30%	—	—	—	—
8. Fully Flexible SEM	L0.8	0	Lisbon	33	42	503064	6480	498961	2018160	5099304	86384.92	86565.22	1910.371	0.71%	—	—	—	—
	S2.8	0	Setúbal	135	10	361020	1056	384697	361020	3977820	2671.5	2775.417	323.45302	0.48%	—	—	—	—

Table 38 - Fleet expansion scenarios computational results.

	Instance Characteristics						Instance Size					Computational Time		Objectives			
							Variables				Non-zero			Coverage		Equity	
Instance	Experiment		Run	City	Demand areas	Stations	Binary	Integer	Continuous	Constraints	coefficients	Root + BC (s)	Total Time (s)	(Z1)	Gap (%)	(Z3)	Gap (%)
L0.3	AEM	+1	0	Lisbon	33	42	212586	4358	213841	866430	2936482	393.94	465.57	1738.973	0.19%	—	—
			2	Lisbon	33	42	212586	4358	213841	866431	3129900	4010.27	4083.2	1738.973	0.19%	16.989	0.00%
		+2	0	Lisbon	33	42	212586	4358	213841	866430	2936482	197.02	267.31	1784.539	0.28%	—	—
			2	Lisbon	33	42	212586	4358	213841	866431	3129900	3881.59	3951.03	1784.539	0.28%	16.989	0.00%
		+3	0	Lisbon	33	42	212586	4358	213841	866430	2936482	159.56	232.87	1822.463	0.50%	—	—
			2	Lisbon	33	42	212586	4358	213841	866431	3129900	1655.91	1725.41	1822.463	0.50%	16.989	0.00%
		+4	0	Lisbon	33	42	212586	4358	213841	866430	2936482	214.05	288.94	1862.397	0.20%	—	—
			2	Lisbon	33	42	212586	4358	213841	866431	3129900	5932.42	6005.22	1862.397	0.20%	16.989	0.00%
	PEM	+1	0	Lisbon	33	42	212586	4358	213841	866430	2936482	96.74	170.03	1740.905	0.16%	—	—
			2	Lisbon	33	42	212586	4358	213841	866431	3129900	582.14	652.15	1740.905	0.16%	16.989	0.00%
		+2	0	Lisbon	33	42	212586	4358	213841	866430	2936482	145.81	219.66	1783.127	0.15%	—	—
			2	Lisbon	33	42	212586	4358	213841	866431	3129900	1359.22	1429.44	1783.127	0.15%	16.989	0.00%
		+3	0	Lisbon	33	42	212586	4358	213841	866430	2936482	140.06	214.7	1819.150	0.37%	—	—
			2	Lisbon	33	42	212586	4358	213841	866431	3129900	1337.91	1407.29	1819.150	0.37%	16.989	0.00%
		+4	0	Lisbon	33	42	212586	4358	213841	866430	2936482	136.59	210.21	1856.460	0.27%	—	—
			2	Lisbon	33	42	212586	4358	213841	866431	3129900	1629.09	1698.63	1856.460	0.27%	16.989	0.00%
	VMER	+1	0	Lisbon	33	42	212586	4358	213841	866430	2936482	21.2	95.94	1713.081	0.43%	—	—
			2	Lisbon	33	42	212586	4358	213841	866431	3129900	120.06	189.45	1713.081	0.43%	16.111	0.00%
		+2	0	Lisbon	33	42	212586	4358	213841	866430	2936482	24.98	98.25	1717.910	0.46%	—	—
			2	Lisbon	33	42	212586	4358	213841	866431	3129900	137.22	207	1717.910	0.46%	16.111	0.00%
		+3	0	Lisbon	33	42	212586	4358	213841	866430	2936482	18.5	62.45	1723.531	0.31%	—	—
			2	Lisbon	33	42	212586	4358	213841	866431	3129900	165.14	234.56	1723.531	0.31%	16.111	0.00%
		+4	0	Lisbon	33	42	212586	4358	213841	866430	2936482	17.11	88.01	1726.585	0.26%	—	—
			2	Lisbon	33	42	212586	4358	213841	866431	3129900	272.22	341.8	1726.585	0.26%	16.111	0.00%
	SIV	+1	0	Lisbon	33	42	212586	4358	213841	866430	2936482	58.58	69.3	1737.669	0.26%	—	—
			2	Lisbon	33	42	212586	4358	213841	866431	3129900	634.31	703.53	1737.669	0.26%	16.111	0.00%
		+2	0	Lisbon	33	42	212586	4358	213841	866430	2936482	40.8	113.03	1786.039	0.24%	—	—
			2	Lisbon	33	42	212586	4358	213841	866431	3129900	798.86	869.41	1786.039	0.24%	16.111	0.00%
		+3	0	Lisbon	33	42	212586	4358	213841	866430	2936482	43.17	113.62	1824.515	0.46%	—	—
			2	Lisbon	33	42	212586	4358	213841	866431	3129900	412.01	481.01	1824.515	0.46%	16.111	0.00%
		+4	0	Lisbon	33	42	212586	4358	213841	866430	2936482	27.86	98.25	1864.177	0.32%	—	—
			2	Lisbon	33	42	212586	4358	213841	866431	3129900	872.91	941.69	1864.177	0.32%	16.111	0.00%
S2.3	AEM	+1	0	Setúbal	135	10	193546	1016	217009	852309	2265540	306.72	361.55	322.8708	0.00%	—	—
			2	Setúbal	135	10	193546	1016	217009	852310	2306592	1210.5	1280.39	322.8708	0.00%	18.233	4.89%
		+2	0	Setúbal	135	10	193546	1016	217009	852309	2265540	248.77	323.66	323.689	0.14%	—	—
			2	Setúbal	135	10	193546	1016	217009	852310	2306592	647.88	721.88	323.689	0.14%	16.686	0.00%
		+3	0	Setúbal	135	10	193546	1016	217009	852309	2265540	250.61	325.18	324.037	0.07%	—	—
			2	Setúbal	135	10	193546	1016	217009	852310	2306592	572.58	647.14	324.037	0.07%	16.686	0.00%
		+4	0	Setúbal	135	10	193546	1016	217009	852309	2265540	256.03	330.31	324.336	0.00%	—	—
			2	Setúbal	135	10	193546	1016	217009	852310	2306592	492.45	566.87	324.336	0.00%	16.686	0.00%
	PEM	+1	0	Setúbal	135	10	193546	1016	217009	852309	2265540	139.93	213.78	320.569	0.00%	—	—
			2	Setúbal	135	10	193546	1016	217009	852310	2306592	136.56	212.69	320.569	0.00%	18.233	0.00%
		+2	0	Setúbal	135	10	193546	1016	217009	852309	2265540	150.3	224.32	320.569	0.00%	—	—
			2	Setúbal	135	10	193546	1016	217009	852310	2306592	145.56	219.75	320.569	0.00%	18.233	0.00%
		+3	0	Setúbal	135	10	193546	1016	217009	852309	2265540	149.59	223.37	320.569	0.00%	—	—
			2	Setúbal	135	10	193546	1016	217009	852310	2306592	150.08	224.49	320.569	0.00%	18.233	0.00%
		+4	0	Setúbal	135	10	193546	1016	217009	852309	2265540	149.19	224.09	320.569	0.00%	—	—

			2	Setúbal	135	10	193546	1016	217009	852310	2306592	146.91	221.45	320.569	0.00%	18.233	0.00%
	VMER	+1	0	Setúbal	135	10	193546	1016	217009	852309	2265540	98.97	172.91	320.549	0.00%	—	—
			2	Setúbal	135	10	193546	1016	217009	852310	2306592	180.63	255.77	320.549	0.00%	18.233	0.00%
		+2	0	Setúbal	135	10	193546	1016	217009	852309	2265540	98.42	172.07	320.549	0.00%	—	—
			2	Setúbal	135	10	193546	1016	217009	852310	2306592	177.72	252.61	320.549	0.00%	18.233	0.00%
		+3	0	Setúbal	135	10	193546	1016	217009	852309	2265540	98.44	171.5	320.549	0.00%	—	—
			2	Setúbal	135	10	193546	1016	217009	852310	2306592	177.23	251.48	320.549	0.00%	18.233	0.00%
		+4	0	Setúbal	135	10	193546	1016	217009	852309	2265540	98.58	172.20	320.549	0.00%	—	—
			2	Setúbal	135	10	193546	1016	217009	852310	2306592	179.70	254.00	320.549	0.00%	18.233	0.00%
	SIV	+1	0	Setúbal	135	10	297514	1251	320581	1267638	3542596	913.23	1053.03	328.625	0.00%	—	—
			2	Setúbal	135	10	297514	1251	320581	1267639	3605593	2219.03	2320.67	328.625	0.00%	18.233	2.53%
		+2	0	Setúbal	135	10	297514	1251	320581	1267638	3542596	452.94	555.11	330.016	0.00%	—	—
			2	Setúbal	135	10	297514	1251	320581	1267639	3605593	1010.05	1111.83	330.016	0.00%	16.686	0.00%
		+3	0	Setúbal	135	10	297514	1251	320581	1267638	3542596	569.52	684.62	330.105	0.02%	—	—
			2	Setúbal	135	10	297514	1251	320581	1267639	3605593	959.34	1061.10	330.105	0.02%	16.686	0.00%
		+4	0	Setúbal	135	10	297514	1251	320581	1267638	3542596	543.34	648.42	330.210	0.00%	—	—
			2	Setúbal	135	10	297514	1251	320581	1267639	3605593	893.44	1013.99	330.210	0.00%	16.686	0.00%
L0.4	AEM	+1	0	Lisbon	33	121	408721	5565	401545	1633448	5581811	1548.3	1640.53	1739.938	0.20%	—	—
			2	Lisbon	33	121	408721	5565	401545	1633449	5953611	77337.55	77466.31	1739.938	0.20%	13.587	4.99%
		+2	0	Lisbon	33	121	408721	5565	401545	1633448	5581811	1708.34	1798.49	1784.982	0.30%	—	—
			2	Lisbon	33	121	408721	5565	401545	1633449	5953611	86406	86533.98	1784.982	0.30%	9.818	9.75%
		+3	0	Lisbon	33	121	408721	5565	401545	1633448	5581811	2428.64	2338.48	1828.082	0.22%	—	—
			2	Lisbon	33	121	408721	5565	401545	1633449	5953611	86403.59	86528.59	1828.082	0.22%	9.818	20.30%
		+4	0	Lisbon	33	121	408721	5565	401545	1633448	5581811	520.75	671.42	1862.676	0.24%	—	—
			2	Lisbon	33	121	408721	5565	401545	1633449	5953611	86552.66	86614.181	1862.676	0.24%	9.818	25.85%
S2.4	AEM	+1	0	Setúbal	135	16	253384	1094	276193	1090212	2962713	711.75	798.9	323.878	0.00%	—	—
			2	Setúbal	135	16	253384	1094	276193	1090213	3018021	3967.13	4054.86	323.878	0.00%	18.233	0.23%
		+2	0	Setúbal	135	16	253384	1094	276193	1090212	2962713	904.47	991.06	324.362	0.00%	—	—
			2	Setúbal	135	16	253384	1094	276193	1090213	3018021	1451.56	1538.67	324.362	0.00%	16.686	0.00%
		+3	0	Setúbal	135	16	253384	1094	276193	1090212	2962713	802.74	890.68	324.738	0.00%	—	—
			2	Setúbal	135	16	253384	1094	276193	1090213	3018021	1550.56	1638.76	324.738	0.00%	16.686	0.00%
		+4	0	Setúbal	135	16	253384	1094	276193	1090212	2962713	778.28	864.84	324.8572	0.00%	—	—
			2	Setúbal	135	16	253384	1094	276193	1090213	3018021	1580.3	1668.49	324.8572	0.00%	16.686	0.00%