Improving routing for medical test sample transportation to a clinical analysis laboratory using optimization approaches

Íris Vieira da Silva

Thesis to obtain the Master of Science Degree in

Biomedical Engineering

Supervisors: Professor Ana Paula Barbosa Póvoa

Professor Tânia Rodrigues Pereira Ramos

Examination Committee

Chairperson: Professor Mónica Duarte Correia de Oliveira

Supervisor: Professor Ana Paula Barbosa Póvoa

Members of the committee: Doctor Daniel Rebelo dos Santos

December 2020

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.

Preface

The work presented in this thesis was performed at SIEMENS Healthineers (Lisbon, Portugal), during the period March-October 2020. The thesis was supervised by Prof. Ana Póvoa and Prof. Tânia Ramos from Instituto Superior Técnico and by Dr. Filipa Matos Baptista from SIEMENS Healthineers.

Acknowledgments

Throughout the writing of this thesis I have received a substantial amount of support and assistance, I take now the opportunity to thank everyone that made it possible.

Firstly, a special note of thankfulness to Professor Ana Póvoa and Professor Tânia Ramos for all the mentoring over the past few months, for always pushing me to go further and do better and for the constant availability. For all the words of encouragement, strength and optimism demonstrated. Without their guidance and persistent help this dissertation would have not been possible.

Secondly, I would like to thank Filipa Baptista for this amazing opportunity. For all the mentoring, attention and advice that I will take for the rest of my life. To the team that also made it possible: Paulina, Mafalda, Carolina, Carlos Parente, Carlos Martins, Vera, Nuno and Manuel.

Thirdly, I wish to express my gratitude to all my friends for always being supportive of me and for the moments of joy they bring to my life. To the ones that I have since forever and to the ones that entered my life in more recent years, I will always be thankful.

Last but certainly not least, to my family. To my parents, Sílvia and Amílcar, for always believing in me and for always being there no matter what. To my brother David, for the encouragement and for all the invaluable support. To Lua, for always being by my side. I will forever be grateful.

Abstract

Biomedical tests take an important role assisting physicians providing accurate diagnostics to patients. These samples are collected in specimen collection centres and are transported in cool boxes to the laboratory where they are analysed. In the present dissertation, the case of a diagnostic provider group is considered. Since this entity deals daily with a great number of samples, the flow of their arrivals must be considered to ensure for a smooth workload in the laboratory. Hence, it is desirable to desynchronize the samples arrivals by handling the routes' schedules and passing orders.

In this work, a Biomedical Sample Transportation Problem (BSTP) mathematical model is developed to assist the entity under study in the creation of their routes. These routes have the particularity that they can depart from different locations other than the laboratory. Further, a multi-start algorithm based on the developed model is introduced with the purpose of handling large instances in a reasonable amount of time. The objective is to both minimize the routes' total travel time and the number of sample boxes' arriving to the laboratory in the busiest time period. To the best of ones' knowledge, none of the previous works on the BSTP has considered the possibility of open routes as the current work presents.

Application of the model in two regions with different objectives suggests that improvements can be achieved. More specifically, for the region that considers both objectives (i.e., minimizing the routes total travel time along with desynchronizing the arrivals) the arrivals are successfully desynchronized, however implying an increase of 5.7% and of 2.6% in the routes' total distance and total time, correspondingly. For the region where the only objective is to minimize the routes' total duration, improvements of 16.2% and of 11.2% can be attained in terms of the total distance and total time, respectively.

Keywords: Vehicle Routing Problem; Biomedical Sample Transportation Problem; Desynchronized arrivals; Open routes; Time-windows; Optimization; Metaheuristic; Multi-start algorithm.

Resumo

As análises clínicas desempenham um papel fundamental a efetuar diagnósticos precisos, por parte dos médicos, aos seus pacientes. Estas amostras são colhidas nos postos de colheita e transportadas em caixas térmicas para o laboratório, onde são analisadas. Na presente dissertação, é considerado o caso particular de um grupo de análises clínicas. Como esta entidade lida diariamente com um elevado número de amostras, deve ter-se em consideração o fluxo das chegadas das mesmas de forma a garantir uma melhor distribuição da carga de trabalho no laboratório. Deste modo, é importante dessincronizar as chegadas das análises clínicas alterando as rotas atualmente estabelecidas.

Assim sendo, é desenvolvido nesta tese um modelo matemático para o Problema de Transporte de Amostras Biomédicas (BSTP) de modo a auxiliar esta entidade na criação das suas rotas. Estas rotas têm a particularidade de poderem partir de diferentes locais para além do laboratório. Para além disso, é sugerida uma heurística *multi-start* baseada no modelo matemático concebido, de modo a ser possível encontrar soluções para instâncias grandes num intervalo de tempo razoável. O propósito é então minimizar o tempo total das rotas e o número de caixas que chegam ao laboratório no período de tempo do dia com maior frequência de chegadas. De facto, nenhum dos trabalhos anteriores sobre o BSTP considerou a possibilidade de rotas abertas como o trabalho atual apresenta.

A aplicação do algoritmo em duas regiões com objetivos diferentes sugere que há melhorias que podem ser alcançadas. Mais especificamente, para a região que considera ambos os objetivos mencionados (ou seja, minimizar o tempo total das rotas e dessincronizar as chegadas) foi possível dessincronizar as chegadas, porém com custo acrescido de 5.7% e de 2.6% na distância total percorrida e do tempo total das rotas, respetivamente. Para a região cujo objetivo é a minimização do tempo total das rotas, puderam ser alcançadas melhorias de 16.2% e 11.2%, respetivamente, em termos da distância total percorrida e do tempo total das rotas.

Palavras-chave: Problema de Roteamento de Veículos; Problema de Transporte de Amostras Biomédicas; Chegadas dessincronizadas; Rotas abertas; Janelas de tempo; Otimização; Metaheurística; Algoritmo *multi-start*.

Contents

List of Figures

List of Tables

Abbreviations

B&B– *Branch-and-bound*

BSCP– *Blood Sample Collection Problem*

BSTP– *Biomedical Sample Transportation Problem*

CP– *Constraint Programming*

CVRP– *Capacitated Vehicle Routing Problem*

ES– *Enterprise Services*

HFVRP– *Heterogeneous Fleet Vehicle Routing Problem*

MILP– *Mixed Integer Linear Program*

MTVRP– *Multi-trip Vehicle Routing Problem*

NP-Hard– *Nondeterministic Polynomial-time Hard*

OCVRP– *Open Capacitated Vehicle Routing Problem*

SCC– *Sample Collection Centre*

SH– *SIEMENS Healthineers*

TSP– *Traveling Salesman Problem*

VRP– *Vehicle Routing Problem*

VRPTW– *Vehicle Routing Problem with Time Windows*

1 Introduction

1.1 Motivation

One of the main challenges addressed by healthcare systems is to guarantee high-quality services having limited resources. Diagnostic laboratories are constantly seeking for ways to improve their capabilities by strengthening diagnostic accuracy, managing higher workloads faster and extend test menus. Simultaneously, they are required to become more efficient by reducing costs [1].

A significant part of a successful laboratory operation is biomedical sample transportation. In fact, a well preserved and timely biomedical sample is vital to lead to an accurate diagnosis. Daily, thousands of biomedical samples are being transported from various healthcare facilities (e.g., hospitals, private clinics and retirement homes) to laboratories to be examined. Even if large hospitals often can locally analyse samples, most centres are not sufficiently equipped to perform such tests; thus, it must be ensured that samples are safely transported to clinical analysis laboratories. Accordingly, many Specimen Collection Centres (SCCs) are assigned to the same laboratory making essential managing the flow of the samples arriving to the laboratory in order to avoid congestion.

Biomedical samples are perishable items; therefore, it must be ensured that once these samples are collected in the SCCs that they arrive in the laboratory in a short amount of time to guarantee that they lead to an accurate diagnosis. This short life span of biomedical samples usually makes necessary to perform more than one visit per day in each SCC in order to avoid sample perishing.

The current study addresses this challenge by exploring the aim of an entity, which pursues its activity in the area of the public health. This entity deals with a great number of biomedical samples each day. Thus, the present research has the objective of improving the routing of biomedical test sample transportation to their principal clinical analysis laboratory using optimization approaches. This improvement should ensure for a smoother workload at the laboratory by desynchronizing the biomedical samples arrivals while at the same time reduce the operational costs and increase the resources usage by minimizing the routes total duration.

1.2 Dissertation objectives

The present dissertation focuses on developing and applying operation research techniques to improve the routes of a group of clinical pathology that transfer biomedical samples from the SCCs in two regions, Region A and Region B, to the central laboratory. The present master thesis makes part of a project from the SIEMENS Healthineers (SH) Enterprise Services (ES) team; thus, the developed work should be in accordance with their principles. SH ES team always seeks at delivering solutions that promote better clinical outcomes, improve patient experience while in parallel reduce costs.

The objective of the present research focuses on developing an optimization model that produces routes that are defined by both the sequence of the SCCs to visit and the schedules at which the SCCs should be visited. The developed model should integrate the objectives defined by the entity for both regions (Region A and Region B). Moreover, notwithstanding the fact that this model is inspired by the entity's case study, it should be general enough to model similar contexts. To achieve this main goal, secondary research goals are also established:

- Reviewing previous research already performed in the field of the Biomedical Sample Transportation Problem (BSTP) and emphasizing their complex nature;
- Identifying potential gaps in the literature that are important for the resolution of the case study and suggesting possible alternative modelling approaches;
- Contributing to the existing literature by providing both an optimization model and a multi-start heuristic for the entity's context;
- Finally, providing possible recommendations to assist the entity in the creation of new routes.

1.3 Proposed methodology

In order to accomplish the objectives, the proposed methodology comprises five fundamental steps, demonstrated in the following figure, Figure 1.1.

Figure 1.1 – Proposed methodology.

Step 1 - Case-study contextualization and problem definition

The aim of the first step is to introduce the entity by describing its activity and the operations related to the transportation of the biomedical samples. Thus, present their network structure, limitations, main challenges faced and the corresponding objectives.

Step 2 - Literature Review

This step focuses on introducing the class of optimization problems found in logistics, the Vehicle Routing Problem (VRP). More specifically, describe the most common mathematical programming formulations used, the VRP variants closer to the case study and the most common resolution methods. Also, surveying the state-of-art of the BSTP and underline the contribution of this thesis to the already existing literature.

Step 3 – Mathematical model formulation and heuristic algorithm presentation

In the third step, the mathematical formulation of the proposed model is outlined, and the structural components are presented. Then, the algorithm proposed to solve the case study is described.

Step 4 - Performance analysis of the solution methods

Afterwards, a performance analysis to validate the solution approach is performed. Therefore, a set of instances are applied both to the developed mathematical model using the optimization software and to the created heuristic algorithm. The corresponding results are then compared to evaluate their performance and to find the parameters that fit best the entity's scope.

Step 5 - Case-study resolution and analysis

Finally, the algorithm is applied to the presented case-study in two different regions in order to produce recommendations. These results are analysed, discussed and compared to the current existing ones.

1.4 Thesis outline

The present dissertation is structured into eight chapters, which are briefly described subsequently.

Chapter 1 – Introduction

The first chapter, the present one, contextualizes the problem by providing a motivation and a general description of the objectives to be attained in this work.

Chapter 2 – Case Study

The aim of the second chapter, the case study, is to characterize the entity, its activity and describe their current biomedical sample collection operations. Namely, it contextualizes and describes the entity's background, outlines the network structure of the SCCs and laboratories and details the current routes and their limitations.

Chapter 3 – Literature Review

After identifying the problem, the literature review aims at understanding the well-known class of optimization problems found in logistics, the VRP. Thereafter, its description, the most common mathematical programming formulations used, the different VRP variants (with more attention to the ones that are closer to the case study) and their resolution methods are described in detail. It also presents the BSTP, the VRP applied to the transportation of biomedical samples, and a description of the studies developed in the past that are closer to the case study, highlighting in particular the contribution of this work to existing studies.

Chapter 4 – The Biomedical Sample Transportation Problem

In this chapter, the mathematical formulation of the proposed linear model is described. Thus, the structural components (i.e., the sets, parameters, decision variables, objective function and the constraints) are presented.

Chapter 5 – Heuristic Algorithm

The type of problems as the one under study hold various characteristics that make it difficult to be solved using exact solution approaches. Thus, the need of developing a heuristic algorithm arises. In Chapter 5, a multi-start algorithm to solve the BSTP with open routes and desynchronized arrivals is theoretically illustrated.

Chapter 6 – Performance Numerical Results

In Chapter 6, the performance of the proposed methods (optimization model and heuristic algorithm) is assessed over many realistic instances based on the network of the Region A of the entity.

Chapter 7 – Case Study Results and Discussion

Chapter 7 comprises the presentation of the results for the real-world case study, applying the developed heuristic algorithm in order to obtain results for both regions.

Chapter 8 – Conclusions and Future Work

The last chapter summarizes the work developed and incorporates the systematization of all the pros and cons of this model. It also presents suggestions for further work.

2 Case Study

The development of the present thesis has as basis a real-world case study and aims at optimizing the biomedical samples collection routes of a medical diagnostic provider group in two regions: Region A and Region B. Thereafter, the present chapter has the objective of characterizing the group, its activity and describing their current biomedical sample collection operations in the referred locations. Hence, it starts with a brief contextualization and description of the entity's background. Then, the associated routes are detailed, and their limitations are listed. Finally, a reflection on the challenge of the thesis is presented.

2.1 Entity's background

The entity under study is a group of clinical pathology which aims at continuously following up both with the medical evolution and challenges in the area of clinical pathology, while providing this knowledge to medical staff and patients. To be able to provide clinical diagnoses, the group has distributed throughout several collection points and laboratories, including hospital laboratories. They also provide sample collections at home, a process valuable for the elderly and for those with reduced mobility, improving patient experience.

An important aspect that lies behind the entity's process of providing medical diagnosis is the associated sample transportation operations and logistics. To be able to provide an accurate medical diagnosis, it is fundamental that the biomedical sample is transferred with appropriate handling from the site where it was taken to the laboratory. The entity has then the concern of coordinating three sources of samples: collections at home, collection points and hospitals.

Typically, patients go directly to the entity's collection points to take the biomedical samples, however sample collections at home can be required and, in this case, professionals must go with the necessary material to the patient's house. Home biomedical sample collections are organized in such a way that it is the responsibility of the nearest collection point to collect the sample and bring it back to the collection point. Thereafter, throughout the day, collection points gather samples collected at patient's home and in these points, which then need to be transferred to the laboratory to be tested.

The entity has additionally an agreement with some hospitals to analyse their samples. Some of these larger hospitals have integrated laboratories that can perform common exams (e.g., haematological tests, standard biochemical tests and urinalysis). This allows some samples not having the need to go to the central laboratory to be tested since they are tested in the hospitals. Nevertheless, when more specific exams are required (e.g., immunological tests and serological tests), samples must be transferred to the central laboratory to be properly examined. Moreover, in the hospitals that do not have an integrated laboratory, all the samples must be transferred to the central laboratory.

Henceforth, with the need of transporting samples from the hospitals/collection points to the central laboratory arises the existence of predefined routes. The aim of the present thesis is to optimize these routes by variating the timing and the order that the collection points and hospitals are visited. Currently, sample collection routes are organized taking into consideration the hospitals and collection points geographical locations and defined collection hours, while ensuring that all the samples in that day arrive to the laboratory and meet all the requirements.

The focus of the present thesis lies on optimizing their routes in Regions A and B that transport biomedical samples to the entity's central laboratory, as this is one of the major challenges that this entity faces.

2.2 Sample collection routes

As biomedical samples are considered perishable items, its transportation must be made considering certain requirements. These include guarantee that the samples arrive to the laboratory with the required quality (for this, cool boxes are used) and that each route has a maximum transportation time restriction, since the samples have a limited time span. Additionally, depending on the opening hours of the collection points/hospitals, more than one route to the same point in the same day can be required.

An important aspect that the entity considers when constructing the routes is the different opening hours of the collection points and hospitals. During weekdays, most collection points are open until late in the morning, allowing the existence of only one collecting route per day in that location. For collection points of bigger dimension and hospitals, since they are open for longer hours, multiple routes can be settled throughout the day, during the morning and afternoon period. During weekends, collection points are only open on Saturday's morning, being closed on Sundays. In the case of hospitals, they continue having the same working hours as on the weekdays.

When constructing the routes, the entity does not consider the vehicles capacity as a limitation. In fact, it has a homogeneous fleet (i.e., the vehicles are all equal) where each vehicle can transport 4 cool boxes. Each one of these cool boxes has the capacity of around 7/8 test tube supports carrying 64 samples each. Therefore, it allows each vehicle to transport a very high number of samples (around 2000) making their capacity not a limitation when collecting the samples.

The demand on each collection point and hospital is not known in advance, since this number fluctuates according to multiple factors (e.g., day of the week, time of the year, etc.). Consequently, the routes are organized in a manner that guarantees that all the samples are transported to the laboratory to be analysed in that day independently of the daily demand on these points. This is possible both due to the unlimited vehicles' capacity and by the multiple collections to the same collection point/hospital.

Regarding the geographical organization of the routes, in the case of Region A, they usually start from the drivers' house and are organized so that they first visit the closest sites. After performing the route, if a new route is attributed to that same driver, the driver then starts from the laboratory. In the case of Region B routes, since these take longer due to the geographic dispersion of the collection points and hospitals, only one route is attributed to each driver. In this case, drivers depart from one of the entity's laboratory located in Region B, go to the determined collection points and hospitals to collect the biomedical samples and transport them to the central laboratory.

It is expected that drivers do not spend more than 5-10 minutes on each point, hence it is the technician's responsibility to have the samples prepared for transportation. In smaller collection points, if the technician perceives that it is not necessary for the driver to pass on that collection point, the driver can be advised and skip it. Besides these predefined routes, the entity is also prepared for any emergency route that can arise (these usually include transference of samples to the laboratory and between laboratories).

Currently, in Region A the entity counts on having 39 fixed established routes that satisfy all the aforementioned restrictions. In this moment, the entity counts on having a total of 28 vehicles available to perform the routes.

Associated with the way the routes in this region are defined arises the issue of having periods of the day where too many samples arrive to the laboratory contrary to others where very few samples arrive. This leads to significant differences in the laboratory throughput throughout the day, that need to be softened. This happens substantially between 2 pm and 4 pm and is, in fact, a bottleneck problem in the samples' supply chain of the entity's central laboratory.

Regarding Region B routes, the entity currently counts on 3 fixed established routes that also satisfy all the restrictions previously presented. Summing to those restrictions, since these routes take several hours to be completed, a lunch break of 1 hour has also to be considered.

The problem of having too many samples arriving in the laboratory in certain periods of the day is not a concern in Region B routes, since these samples usually arrive to the laboratory in the evening. Nevertheless, they still have the issue of taking a long time to be completed, which is a challenge, given the perishable nature of the biomedical samples transferred.

As a consequence of the challenges described for both these regions, the focus of the present thesis is to address these issues by constructing and proposing routes that, on one hand, allow for a smoother workload at the laboratory, and on the other hand, reduce the operational costs and increase the resources usage by minimizing the routes total duration.

2.3 Chapter conclusions

With the understanding of the logistics and operations associated with the entity biomedical sample collection, one can conclude that the design of the collection routes is a highly complex problem. Not only it is necessary to attend their restrictions and limitations, but also it is important to take their objectives into consideration. In fact, the biggest challenge is to face the bottleneck in the samples supply chain related with the samples' arrival time while, at the same time, make the routes efficient by trying to reduce transportation costs by minimizing the routes' duration.

It is, therefore, the aim of this work to address this challenge by redesigning and rescheduling the collection routes in the two regions under study, considering the specific needs and challenges of each of them. More specifically, for Region A it is important to tackle the challenge of desynchronizing the samples' arrivals in the laboratory, while at the same time take into concern the reduction of transportation costs; for Region B the main concern is to minimize the routes' total duration. In the next chapter, a literature review on the issue is explored to perceive how to address the problem in hand and face these challenges.

3 Literature Review

The present chapter has the objective of understanding deeply the well-known class of optimization problems found in logistics, the Vehicle Routing Problem (VRP), that can be applied to solve the problem previously presented in the Case Study. Thereafter, its description, the most common mathematical programming formulations used, the different VRP variants (with more attention to the ones that are closer to the case study) and their resolution methods are presented in detail. Afterwards, the Biomedical Sample Transportation Problem (BSTP) is presented, and a description of the studies developed in the past that are closer to the case study are reviewed.

3.1 Vehicle Routing Problem

The VRP constitutes a well-known class of optimization problems found in logistics, originated from real-life needs [2]. The problem was first introduced in 1959 by Dantzig and Ramser [3] as a generalization of the Traveling Salesman Problem (TSP) presented by Flood, in 1956 [4]. The TSP arises from the situation of a salesman who wants to visit his clients in a given set of cities and then return to his own city, in the shortest possible route.

In the classical VRP, a fleet of vehicles is based at a single depot to collect or deliver products for a set of geographically scattered nodes [5]. Each of these vehicles, whose capacity cannot be exceeded, leaves the depot, visits the nodes and returns to the depot. The aim is to find the optimal set of routes so that all the nodes are served exactly once. Thereafter, fundamental elements of a typical VRP include the road network, the vehicles, the nodes and the depot [6]. Moreover, depending on the problem, the optimization objective can vary. Typical objectives are the minimization of the overall transportation costs which depends on the global travelled distance and on the fixed costs associated with the vehicles, the minimization of the number of vehicles needed to serve the nodes and the minimization of the penalties concerning partial service of the nodes [7]. Figure 3.1 represents a typical VRP solution.

In the early 2000s, most methodological studies were based on a limited subset of operational problems [8], thereafter, on account of real-life needs, literature on routing problems has widely expanded with several more attributes/variants being introduced in the problem and, with faster and more efficient solution methods being developed [9].

The optimal solution to the VRP is considered Nondeterministic Polynomial-time Hard (NP-Hard), as a result, its solvability using mathematical programming is conditioned and only small instances can be solved optimally [10]. Thus, to solve moderate-size problems that are of interest in practice, heuristics are often used [11].

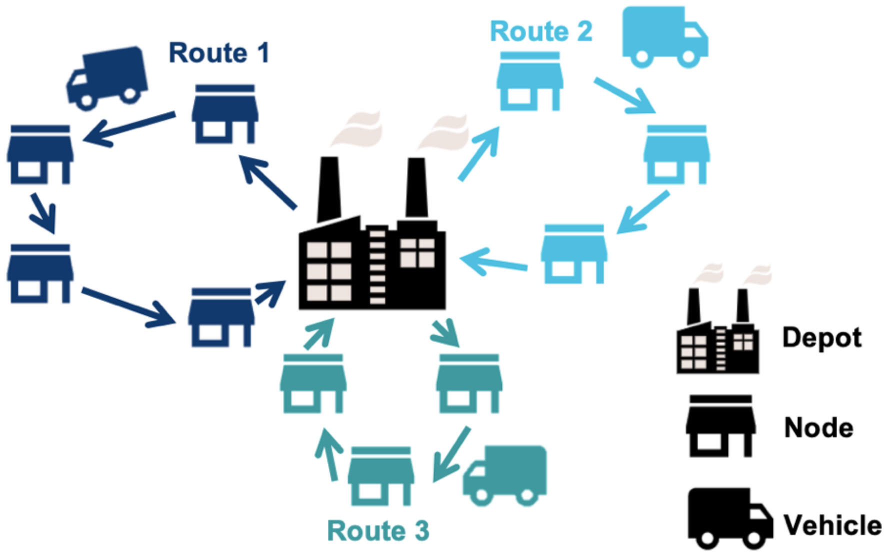


Figure 3.1 – Example of routes elaborated by solving a classical VRP (adapted from [12]).

3.2 Basic models for the VRP

Typically, to represent the road network of a VRP, a graph, *G*, is used. *G* is composed by a set of edges, *A*, that represent the road sections and by a set of vertices, *V*, that represent the depot and the nodes, so that $G = \{A, V\}$ [7]. The vertex set, $V = \{0, 1, ..., n\}$, is usually formulated in such manner that the depot corresponds to vertex *0* and the rest of the vertices *1, ..., n* correspond to the set of nodes that are needed to be visited in the problem. On its turn, the edge set is commonly represented by $A = \{(i, j): i, j \in V, i \neq j\}$, where *i* and *j* represent the limits of the section. Many mathematical programming formulations for the VRP have been suggested in the literature though, the most common formulations include the vehicle flow formulation, the commodity flow formulation and the set partitioning formulation [7].

Vehicle flow formulation models use variables for each edge of the graph [13], which include the two-index vehicle flow formulation and the three-index vehicle flow formulation. The two-index vehicle flow formulation uses the integer variable x_{ij}, which represents the number of times edge *(i,j)* appears in the optimal solution. The three-index vehicle flow formulation, on the other hand, uses the binary variable x_{ijk} to represent the routing solution: x_{ijk} takes the value of 1 if edge *(i,j)* is traversed by vehicle *k* and takes the value 0 otherwise [14].

In commodity flow formulation models, integer variables related with the edges are added to the two- or three-index flow formulations [7]. More specifically, the two-commodity flow formulation uses the binary variable x_{ij}, which takes the value of 1 if the edge *(i,j)* is in the solution, and uses the flow variables y_{ij} and y_{ji} that represent, respectively, the load and the empty space on the vehicle when the edge *(i,j)* is traversed [14].

Finally, models of set partitioning formulations type have an exponential number of binary variables that are associated with different feasible vehicle routes [7]. In fact, this formulation uses the binary variable

x_j that is associated with each feasible route *j*, which takes the value of 1, if the route *j* is in the optimal solution [14].

3.3 VRP variants

Depending on the operational context, the problem can have different constraints, designed to represent the problems characteristics, introduced in each of the fundamental elements (i.e., road network, vehicles, nodes and depots) that leads to the generation of several VRP variants. Typical constraints include time windows, multiple depots, multiple periods, and heterogeneous fleets [15], [16].

Since both characteristics and restrictions of the problem treated in the present thesis were previously presented in the case study section, the VRP variants that are more relevant for this purpose are now be reviewed in detail. These are: the Capacitated Vehicle Routing Problem (CVRP), the Multi-Trip Vehicle Routing Problem (MTVRP) and the Vehicle Routing Problem with Time-Windows (VRPTW).

3.3.1 Capacitated Vehicle Routing Problem (CVRP)

The Capacitated Vehicle Routing Problem is considered the simplest variation of the VRP. Comparing with the TSP, it allows the use of more than one vehicle and assumes that the vehicles' capacity is limited [17]. Thereafter, in the CVRP a fleet of identical vehicles, a homogeneous fleet, positioned at a central depot has to be optimally routed to supply a set of nodes [6], [18]. The demand of each node is known in advance; therefore, the solution has to be built so that the vehicles that leave the depot visit their respective nodes and return to the same depot while guaranteeing that their capacity is not exceeded and that all the nodes are served [6].

As any other classical VRP, the CVRP can be modelled so that it matches other problems characteristics. Nevertheless, there are certain restrictions that, if applied, can lead to alterations of the problem, originating new VRP classes. In fact, a well-known extension of the CVRP lies on when the fleet of vehicles is heterogeneous, Heterogeneous Fleet Vehicle Routing Problem (HFVRP). This heterogenicity can arise from the vehicles' variability in the fixed and operational requirements, including capacity, equipment, speed and cost [19]. Another example of a CVRP extension lies on the Open Capacitated Vehicle Routing Problem (OCVRP). In this variant instead of having routes that begin and end at the depot (i.e., closed routes), as it occurs in the classical CVRP, the routes must either start or end at the depot, but not both at the same time (i.e., open routes) [20].

3.3.2 Multi-trip Vehicle Routing Problem (MTVRP)

A general assumption of the classical VRP is that each individual vehicle only performs one route over the planning horizon [21]. This assumption can be impractical, for instance, in scenarios where the fleet of vehicles is constituted by small vehicles with low capacity, since they do not have the capability of visiting a large number of nodes in each route [22], a common situation faced in metropolitan cities.

The Multi-trip Vehicle Routing Problem emerges as an extension of the classical VRP that allows each vehicle to perform more than one route within a given time period. This problem considers dependency between routes, since it requires that the temporal aspects of the vehicle movements are considered. In fact, it is explicitly needed to guarantee that overlapping routes for the same vehicle do not occur and that the vehicles' total duration is less than or equal to the maximum driving time. Hence, the MTVRP aims at designing a set of schedules that minimizes the total costs in such way that all the nodes are visited [23].

3.3.3 Vehicle Routing Problem with Time Windows (VRPTW)

Finally, another important variant of the VRP is the Vehicle Routing Problem with Time Windows. This variant arises when the service of each node i is limited to a specific time window, [a_i, b_i], requiring the determination of a schedule for the vehicle routes', so that the nodes are served in their corresponding time windows [24].

Depending on the characteristics of the problem, two different types of time windows can arise: hard and soft time windows [25]. If the time window is hard it means that all the nodes must strictly be served within their respective time window. Therefore, the vehicles are not allowed to arrive after the latest time to begin service, b_i, and if they arrive before the node is ready to begin service, a_i, they must wait. On the other hand, if the time window is soft it means that the time frame is allowed to be violated but with penalty cost. Moreover, the VRPTW is usually solved considering two objective functions: minimizing the number of vehicles and minimizing the total travel time [26].

The VRPTW has several useful real-world applications for problems where the nodes have limited time stamps. Some of these examples include bank deliveries, postal deliveries and school bus routing [27].

3.4 VRP solution methods

With the purpose of deciding the best solution method or algorithm to solve a VRP, one must consider both the solution quality and the time taken to obtain it [28]. Since the 1980s, many approaches including exact methods and heuristics have been designed to solve the VRP. Even though exact methods can solve optimally small instances, up to around 100 nodes [29], when facing larger and real-life scenarios, these methods become unfeasible given the time required to obtain the solution. Hence, the use of heuristic procedures in these environments is more appropriate. Although heuristic methods may not ensure an optimal solution, they allow reasonable time solving [29]. The next subsections describe in detail each of the aforementioned approaches.

3.4.1 Exact algorithms

Exact approaches allow to find the optimal solution, while satisfying all the restrictions required by the problem. Nevertheless, due to the high complexity that real-life problems usually have, it may not be possible to find the optimal solution within reasonable time. As new inputs are introduced in the problem, the time of response of the model grows exponentially thus, most times, these methods are only viable for smaller problems. According to Blocho (2020a), the most relevant methods to find the optimal solution are Integer Linear Programming (ILP) and Constraint Programming (CP).

Integer Linear Programming

When the problems are formulated as ILP, the variables are defined as integers (e.g., the vehicles' capacity, constraints associated with the time windows) and the constraints are expressed as inequalities or linear equations. A classical approach to solve the ILP is the well-known branch-and-bound (B&B) method developed by Land and Doig, in 1960 [30]. This technique is based on making partitions of the complete problem obtaining subproblems of smaller dimension that are easier to solve than the complete original problem. The unexplored subproblems generate other new subproblems by dividing the solution space into smaller regions, creating a tree structure which can be solved recursively (i.e., branching). At the same time, rules are used to eliminate regions of the search space that are probably suboptimal (i.e., bounding). When the whole tree has been examined, the best solution in the search is then returned [31]. An example of a B&B structure is represented in Figure 3.2. More examples of strategies widely used in solving the ILP are the branch-and-cut, branch-and-price and branch-and-cut-and-price methods [32].

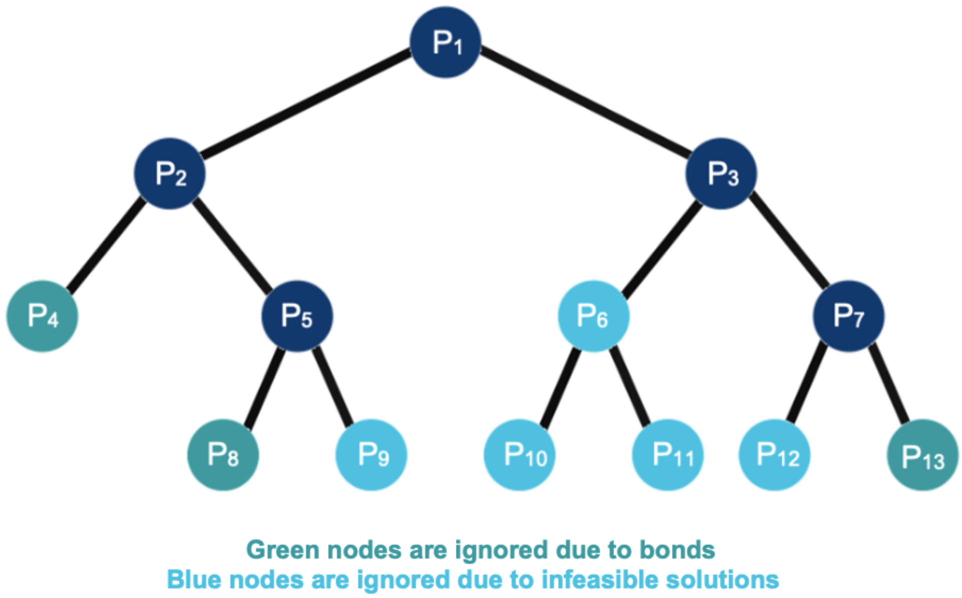


Figure 3.2 – Illustration of the B&B method (adapted from [32]).

Constraint Programming

In CP, a methodology used to represent and solve several optimization problems, the problems are articulated regarding the variables and constraints between the variables and are then solved using complete search techniques. The constraints determine the properties of the optimal solution, not the sequence of steps to be performed. CP is therefore frequently used for several variants of the VRP since these can be stated in terms of capacity, time windows, service time, precedence constraints, etc [32].

3.4.2 Heuristics

A heuristic is a computational method designed to solve larger problems, known to be complicated to solve optimally [27]. Thereby, heuristics allow to obtain feasible, but not necessarily optimal solutions, in reasonable computer time.

According to Laporte (2007), heuristics are subdivided into classical heuristics and metaheuristics. Classical heuristics can additionally be divided into constructive heuristics and improvement heuristics, where both methods execute a limited examination of the solution space. Metaheuristics, on the other hand, are an improvement of classical heuristics with better exploration of the solution space [33]. The referred computational methods are now further explained.

Constructive heuristics

Constructive heuristics start from an empty solution that iteratively extends the current solution until the final one is built. These methodologies usually do not consider any improvement stages. According to Laporte (2007), the most popular constructive heuristic is the savings algorithm developed by Clarke and Wright [34]. Its main objective is to minimize costs by combining routes. Initially, the number of routes equals the number of nodes, as each individual route is constituted by the vehicle leaving the depot, passing in one node and then returning to the depot. A general iteration comprises the connection of two routes, one ending at node *i* and another route beginning at node *j*, resulting in a feasible route *(0, i, j, 0)*, where *0* represents the depot. The saving between nodes *i* and *j*, $Saving_{ij}$, is calculated using the following equation:

$$Saving_{ij} = c_{i0} + c_{0j} - c_{ij} \tag{3.1}$$

, where, c_{ij} represents the cost of travelling from node *i* to node *j*. At each iteration each possible saving is calculated and the one with the greatest saving is selected to be merged. The solution is returned when no more profitable and feasible merges can be found.

The nearest-neighbour and insertion algorithms are other construction heuristics widely used to solve VRPs [35]. The nearest-neighbour method, proposed by Flood in 1956 [4], consists in constructing the solution with a greedy algorithm by selecting the closest unserved node each time. As for insertion heuristics, the generation of feasible solutions is made by inserting unserved nodes into partial feasible routes.

Improvement heuristics

Improvement heuristic algorithms aim at improving any feasible solution by conducting node exchanges between or within vehicle routes. Consequently, they can be separated into two types: intra-route and inter-route heuristics.

Intra-route methods conduct node exchanges within vehicle routes, post optimizing each route separately. Thus, since each route is considered individually it leads to the possibility of applying TSP improvement heuristics due to the fact that TSP considers only single routes [36]. In fact, most of these improvement procedures can be described in terms of the λ-opt algorithm of Lin (1965). In this case, λ nodes are removed from the route, which creates λ disconnected segments. Then, the λ nodes are added and the λ segments are reconnected so that a feasible route is reconstructed, if any profitable reconnection is found it is implemented [37]. This procedure stops when no more improvements can be obtained.

Inter-route methods, on the other hand, consider different routes at the same time, and the objective is to move nodes between them. In this case, procedures that utilize the multi-route structure can be developed. Thompson *et al.* (1993) [38], Breedam *et al.* (1994) [39] and Kindervater *et al.* (2003) [40] provide descriptions of multi-route node exchanges used by several authors [41].

Moreover, according to Laporte (2007) the performance of classical improvement heuristics is good, but not excellent. Thus, it makes it more used as building blocks in metaheuristic methods.

Metaheuristics

Metaheuristics perform a more comprehensive search of the solution space to avoid converging into the first local optimum encountered. Therefore, they commonly combine neighbourhood search rules and recombination of solutions [35]. The neighbourhood consists in all solutions that can be reached by applying a given type of transformation in the solution. When comparing with classical heuristics, metaheuristics provide better quality solutions, however, require more computing time. Moreover, these techniques are dependent on the context of the problem which can make them difficult to extend to different contexts. Overall, the metaheuristics algorithms that have been proposed to solve the VRP can be composed by local search, population search and learning mechanism groups; nevertheless, the greatest algorithms derive from a combination of these different methods [35].

A local search heuristic usually begins with an initial solution, which may be infeasible, and at each iteration moves from the present solution to another one located in the neighbourhood. The search ends with the best-known solution after a stopping criterion has been satisfied [33]. An example of a well-known local search metaheuristic is the Tabu Search introduced in 1986 by Glover [42]. Moreover, some search methods based on local optimization intend to achieve global optima (i.e., to find the optimal value among all possible solutions and not only the one found in the within the neighbourhood set of the candidate solution), thus they require some sort of randomization in order to overcome local optimality. If this diversification is not considered, these methods can get stuck in a small area of the solution space and may not be able to find a global optimum. One strategy to overcome local optimality lies in re-starting the search from a new solution once a region has been widely explored, these methods are called Multi-Start Methods [43].

Population search methods essentially rely on genetic algorithms. Genetic algorithms are based on the evolution process found in nature, as they mimic the process of natural selection. Therefore, they consist on establishing generations based on a population, where each generation is created by selection, recombination and mutation of all the population solutions [35]. A properly designed genetic algorithm should maintain right balance between quality and diversity in a population in order to support efficient search.

Finally, learning algorithms are essentially composed of neural networks and colony optimization algorithms [36]. Neural networks are models established by interconnected units with weighted links mimicking neurons in the brain, that allow to build a solution with the use of feedback mechanisms. In fact, the weights in the networks are gradually modified to better correspond an observed output to a described output. Elastic net and self-organizing maps are VRP applications that use neural networks and one can find examples of their use in Ouduour *et al.* (2020) [44] and Yoshiike *et al.* (2002) [45]. On the other hand, colony algorithms are based in the behaviour of ant colonies found in nature. Ants, when foraging for food release pheromones on their trail, as a result, as time passes more pheromones are released in the most frequent paths. It has been demonstrated experimentally that this ant colony pheromone trail behaviour can lead to the emergence of shortest paths. Similarly, when constructing a VRP solution to a certain move, the probability of it being selected can increase, if in the past this particular move has led to better outcomes [26], [35].

3.5 Biomedical Sample Transportation Problem

Now that all the relevant VRP classes and solution methodologies have been reviewed in detail, the BSTP can be introduced. The BSTP is a challenging VRP arising in the context of healthcare logistics and that can be applied specifically to the case study of the present work. This problem aims at creating a transportation plan to pick up perishable items, biomedical samples, at given locations, referred to as SCCs and to take them to facilities that have adequate treatment equipment, usually central laboratories. Thus, the transportation network consists of a group of SCCs (nodes) affected to a laboratory (depot).

There are some concerns that arise with the biomedical sample transportation, following described. After collection, the samples are consolidated in cool boxes that need to arrive at the laboratory within a given time frame to be treated otherwise, the samples deteriorate and become unusable, increasing the laboratory’s costs and decreasing the quality of the service. In order to respect the sample lifespans, SCCs cannot maintain the collected biomedical samples for a long time; thereby, each SCC can have different number of sample transportation requests according to the SCC open hours, causing an interdependence of pickup times and routes. It is also important to impose a time constraint on the duration of each route to guarantee the samples’ quality. Thus, the BSTP is typically characterized by multiple visits to each node, a time-window on each visit time and multiple routes for each vehicle [46], being closely related to the MTVRP and the VRPTW elucidated previously in section 3.3.2 and 3.3.3, respectively.

All the aforementioned constraints fit the scope of the case study presented in this thesis though, there is also the need of considering desynchronization of the samples' arrival to the laboratory and the possibility of open routes, since that the first route of each vehicle can start in a different location than the laboratory.

Works similar to the case study

Prior to the arrival of the BSTP, most studies in the area of transportation of medicine related perishable items were focused on the specific case of blood transportation, also known as the Blood Sample Collection Problem (BSCP).

In fact, one of the first works to address the problem of blood transportation was developed by Yi in 2003 and was applied to the case of the American Red Cross. The developed model was based on a variant of the VRPTW, but with only one pickup for each centre; their goal was to maximize the quantity of collected and treated blood by minimizing the transportation costs, and, for this end, an exact algorithm was used [47].

Then, Doerner *et al.* (2008) studied the blood transportation course of the Austrian Red Cross, with the particularity that the sample collection was done at mobile collection sites [48]. In this study, it was considered that the moment that the blood deterioration process starts is immediately after the donation, which made them have to consider interdependent pickups.

Years later, in 2010, Ghandforoush *et al.* proposed a decision support system for platelet production and blood mobile scheduling. In their work, the authors only considered single route visiting, in other words, each route consisted in collecting the platelets form a single bloodmobile and returning to the laboratory [49].

Mobasher *et al.* (2015), then proposed a solution to a similar problem, but also with the integration of scheduling the appointments in collection centres. They solved this problem using both an exact approach and an heuristic [50]. Also in 2015, Şahinyazan *et al.* presented a tour mobile collection system for the Red Cross in Turkey, where the main concern was to define tours for mobile collection units and shuttles to collect the harvested blood [51]. Nevertheless, a time constraint for returning to the depot was not considered.

Regarding the BSTP, Anaya-Arenas *et al.* (2016) presented a version of this problem applied to a real-world case study in the Province of Québec in Canada. In this work, each SCC required several collection visits and each one of them needed to happen inside a given and independent time windows that satisfied the samples' lifespan. Therefore, with the aim of minimizing the route duration time a multi-start heuristic was proposed [52].

Using the same context, years later, in 2018, Toschi *et al.*, developed a metaheuristic to solve the BSTP that considered the SCC opening and closing hours as decision variables and, also the interdependency between routes. The objective focused on minimizing the routes' total duration [53].

Last year, in 2019, Anaya-Arenas proposed another metaheuristic to solve the BSTP considering the opening and closing hours of the SCCs and the moment they are visited as decision variables, while taking into consideration the interdependency between routes [54]. Nevertheless, none of the presented works considered the need for the desynchronization of the samples' arrival to the laboratory.

Closer to the context of the case study presented in this thesis, Naji-Azimi *et al.* (2016) studied the BSTP with the desynchronization of the vehicles' arrivals to the depot. To the best of ones' knowledge, they were the first and the only ones yet to consider the desynchronization of the vehicles in this context as an objective. This consideration is extremely important, due to the fact that if too many boxes of samples arrive in a short period of time to the central laboratory, samples are queued and may have to wait a long time before they are analysed, creating a bottleneck in the samples' supply chain [55]. Thus, by minimizing the maximum amount of sample boxes arriving within a given time period it normalizes both the laboratory workload and reduces sample losses. Nevertheless, in Naji-Azimi *et al.* (2016) work the possibility of open routes was not considered.

3.6 Chapter conclusions

The literature review allowed to perceive that one cannot always apply a single specific VRP variant to represent a real-world problem since its complexity usually requires a mixture of these, an example of this lies on the case of the BSTP. Apparently, the most suitable approach to model the case study context and their objective seems to be the model developed by Naji-Azimi *et al.* (2016), where desynchronized arrivals to the depot are considered, one of the main focuses of the current study. Furthermore, it also considers that each SCC has its own time-window, which corresponds to the same situation as the one presented in the case study. Moreover, in this same study, the fleet of vehicles is homogeneous, and the vehicles capacity is considered unlimited, another similarity with the situation of the situation presented in the case study. Regarding the route's duration, the model proposed by Naji-Azimi *et al.* (2016) also considers that it must have a maximum duration time to guarantee the samples' quality, a significant requirement.

Even though Naji-Azimi *et al.* practices truly fit the purpose of the current study, there are certain disparities to take under concern. These includes the fact that open routes were not considered and have to in the case study previously presented, as in some routes the vehicles depart from the drivers' houses or from another laboratory and not from the depot.

Thereafter, in the next chapter a mathematical model, based on this work, but considering constraints lacking, is constructed aiming at simulating the case study reality, while considering the concerns and objectives of the management.

4 The Biomedical Sample Transportation Model

As mentioned throughout this work, the main purpose of this thesis is to assist a clinical provider group improving the planning of their biomedical samples' transportation routes by constructing a model capable of designing routes that determine the visiting sequences of the SCCs and the corresponding schedules. The developed model has the aim of minimizing the routes' duration and the maximum number of samples' arriving at the laboratory within the busiest time period, in order to allow for a smoother workload at the laboratory.

Thereafter, this chapter comprises the mathematical model developed to address and simulate their situation. It is, in fact, an extension of the work found in Naji-Azimi *et al.* (2016). The problem is firstly described in a formal manner in section 4.1. Then, in section 4.2, its mathematical formulation is presented with the description of all the structural elements of the developed model, more specifically, the sets, the parameters and the decision variables, followed by the objective function and problems' constraints. Finally, section 4.3 presents the conclusions of the present chapter.

4.1 Formalization of the problem

In order to formulate the BSTP with desynchronized arrivals, it is firstly important to distinguish both the SCCs locations and the collection requests, since each SCC can have more than one collection request. The n SCCs are defined as:

$$V' = \{v_1', \ldots, v_n'\} \tag{4.1}$$

where each SCC l requires Q_l collection requests leading to a total of $p = \sum_{l=1}^{n} Q_l$ requests. Each one of these requests is composed of one box that contains several biomedical samples.

The BSTP can be modelled as a complete graph $G = \{V, A\}$, where:

$$V = \{v_0, v_1, \ldots, v_p, v_{p+1}, v_{p+2}, \ldots, v_{p+d+1}\} \tag{4.2}$$

is the set of nodes in the network. p represents the number of transportation requests and d the number of drivers. More specifically, the subset $P = \{v_1, v_2, \ldots, v_p\}$ corresponds to the transportation request nodes. Additionally, P_l is denoted as the set of request nodes found in V which matches the same SCC location. Since the first route of each day of each vehicle may depart from a different location rather than the central laboratory (e.g., from the drivers' house in the case of Region A and from the entity's other lab in the case of Region B), the departing places must also be presented in the set of nodes in the network, therefore they are represented by the nodes $\{v_{p+2}, \ldots, v_{p+d+1}\}$. Finally, the laboratory is represented by the nodes $\{v_0, v_{p+1}\}$. Except for the first route of each day, which may start in a different location and finish in the lab, all of the other routes start and finish in the laboratory.

The arc set of the graph can then be defined as:

$$A = \{(v_i, v_j) : v_i, v_j \in V, \quad i \neq j, \quad i,j = 0, \dots, p+d+1 \; and \; not \; (i = 0 \; V \; i = p+1; \; j = p+2, \dots, p+d+1\} \tag{4.3}$$

where a travel time between nodes, t_{ij}, is assigned to each arc (v_i,v_j). Moreover:

$$t_{ij} = 0 \quad if \quad \exists \; l \text{ such that } i, \; j \; \in P_l \tag{4.4}$$

, in other words, if i and j correspond to two requests from the same SCC the travel time to go from one to the other is zero. It is important to note that in general $t_{ij} \neq t_{ji}$, being one of the reasons why a directed graph is used to model the problem. Furthermore, each collection request j must be carried out within a given time window $[a_j, b_j]$. Finally, two requests of the same SCC cannot be served in the same route.

Further, to carry out the transportation requests there are available K uncapacitated vehicles. Each one of these can execute multiple routes ($r = 1, \dots, R$). Each vehicle has a limitation on the length on the working day, T_k. Additionally, both the average loading time for each transportation request in the SCCs, τ_i, and the unloading time at the laboratory before a new route starts, τ_0, must be considered. Moreover, as mentioned in previous sections, the perishable nature of the biomedical samples requires having a maximal transportation time associated with the requests. Thereafter, there is also a maximal transportation time for the samples of request i, T_{max}^i.

The objective of this problem is then to minimize both the routes' total duration (including the waiting times at the SCCs) and a weighted penalty θ related with the maximum number of boxes that arrive in the laboratory in the busiest time period.

4.2 Mathematical formulation

In this section the details of the mathematical model used to define the BSTP with desynchronized arrivals and open routes are presented. Therefore, the sets, the parameters, the decision variables, the objective function and the constraint equations are presented below.

Sets

i – index of the nodes: $i = 0, \dots, p+d+1$;

j – index of the nodes: $j = 0, \dots, p+d+1$;

k – vehicles: $k = 0, \dots, K$;

r – routes per vehicle: $r = 0, \dots, R$;

t – time periods: $t = 0, \dots, T$.

Parameters

T_k – limit on the working day of vehicle k;

τ_i – loading time of request i;

τ_0 – unloading time at the laboratory;

T_{max}^{i} – maximal transportation time for the samples associated with request i;

θ – weighted penalty linked to the maximum number of boxes that arrive at the laboratory during the busiest time period;

d – number of drivers;

a_i – start of time windows of request i;

b_i – end of time windows of request i;

t_{ij} – travel time from node i to node j;

ω – units to discretize time;

M – large constant.

Decision variables

x_{ijkr} $\{0,1\}$ takes the value of 1 if vehicle k in its route r travels from request i to request j; takes the value of 0 otherwise;

y_{itkr} $\{0,1\}$ takes the value of 1 if request i carried out by vehicle k in its route r arrives at the laboratory during the t^{th} time period; takes the value of 0 otherwise;

u_{ikr} R_0^+ represents the visit time (start of loading) of the r^{th} route of vehicle k of the transportation request i;

w N_0^+ represents the highest number of boxes arriving to the laboratory during the most visited time period.

Objective function

The objective of the problem is to desynchronize the biomedical sample arrivals in the laboratory while trying to minimize the total costs, thus the objective function can be modelled as:

$$Min \quad \sum_{k=1}^{K} \sum_{r=1}^{R} \left(u_{p+1kr} - u_{0kr}\right) + \theta \cdot w \tag{4.5}$$

where the first part of the objective function minimizes the routes' total duration (i.e., the sum of the end time minus starting time of each route of each vehicle) which includes the necessary waiting time at each SCC. The second part of the equation, on the other hand, implements a penalty factor θ to the maximum number of boxes arriving during the busiest time period, w.

Constraints

In order to formulate the model so that the solution satisfies all the characteristics previously described in section 4.1, several constraints must be defined. To fully understand them it is firstly important to comprehend that since the first route of each day may not depart from the central laboratory some constraints with the same purpose had to be divided into two different equations: one regarding the first route of each vehicle (i.e., when r equals 1) and another for the other routes (i.e., when r is greater than 1). All the constraints are now presented:

1) All collection requests must be served exactly once

Constraints 4.6 assures that each collection request is serviced by exactly one route.

$$\sum_{k=1}^{K}\sum_{r=1}^{R}\sum_{i=0\ \&\ i\neq p+1}^{p+d+1} x_{ijkr} = 1 \qquad j = 1, \dots, p \tag{4.6}$$

2) First route of each vehicle departs from the drivers' houses

The following constraints, constraints 4.7, ensure that the first route of each day (i.e., when r equals 1) departs from one of the possible departing places that is not the central laboratory.

$$\sum_{k=1}^{K}\sum_{j=1}^{p} x_{ijkr} = 1 \qquad i = p+2, \dots, p+d+1;\ \ r = 1 \tag{4.7}$$

3) Each route services one original point at a time

Each route should only visit one original point at a time:

$$\sum_{j\in P_l}\sum_{i=0\ \&\ i\neq p+1}^{p+d+1} x_{ijkr} \leq 1 \qquad l = 1, \dots, n;\ \ k = 1, \dots, K;\ \ r = 1, \dots, R \tag{4.8}$$

4) Flow conservation

Flow conservation is ensured by constraints 4.9 and 4.10. Constraints 4.9 ensures flow conservation for the routes that start in the lab (i.e., for $r > 1$) and constraints 4.10 for the first route of each vehicle (i.e., for $r = 1$).

$$\sum_{i=0}^{p} x_{ijkr} - \sum_{l=1}^{p+1} x_{jlkr} = 0 \qquad r = 2, \dots, R;\ \ j = 1, \dots, p;\ \ k = 1, \dots, K \tag{4.9}$$

$$\sum_{i=1\&\ i\neq p+1}^{p+d+1} x_{ijkr} - \sum_{l=1}^{p+1} x_{jlkr} = 0 \qquad r = 1;\ \ j = 1, \dots, p;\ \ k = 1, \dots, K \tag{4.10}$$

5) Vehicles can start a route or not

The following constraints state that the vehicle k can start a route r or not. Constraints 4.11 refer to routes that start in the laboratory and constraints 4.12 to the ones that start in a different location.

$$\sum_{j=1}^{p} x_{0jkr} \leq 1 \qquad r = 2, \dots, R;\ \ k = 1, \dots, K \qquad (4.11)$$

$$\sum_{j=1}^{p} x_{ijkr} \leq 1 \qquad r = 1;\ \ k = 1, \dots, K;\ \ i = p+2, \dots, p+d+1 \qquad (4.12)$$

6) Ensures that all routes finish in the laboratory

If a given route has started it has to finish in the laboratory to deliver the samples. Thereafter, constraints 4.13 and 4.14 guarantee it for both routes that start in the laboratory (constraints 4.13) and the ones that start in a different location (constraints 4.14).

$$\sum_{j=1}^{p} x_{0jkr} - \sum_{j=1}^{p} x_{j,p+1kr} = 0 \qquad r = 2, \dots, R;\ \ k = 1, \dots, K \qquad (4.13)$$

$$\sum_{i=p+2}^{p+d+1} \sum_{j=1}^{p} x_{ijkr} - \sum_{j=1}^{p} x_{j,p+1kr} = 0 \qquad r = 1;\ \ k = 1, \dots, K;\ \ i = p+2, \dots, p+d+1 \qquad (4.14)$$

7) Time windows constraints

Constraints 4.15 states that the time windows for each request must be respected.

$$a_j - T_k \cdot \left(1 - \sum_{i=0\&i\neq p+1}^{p+d+1} x_{ijkr}\right) \leq u_{jkr} \leq b_j + T_k \cdot \left(1 - \sum_{i=0\&i\neq p+1}^{p+d+1} x_{ijkr}\right) \qquad \begin{array}{c} j = 0, \dots, p+d+1; \\ k = 1, \dots, K; \\ r = 1, \dots, R \end{array} \qquad (4.15)$$

8) Sub-tour elimination constraints

The sub-tour elimination constraints are represented by the constraints 4.16 and 4.17, once again, separated based on the departing location. Constraints 4.16 characterize the routes that start in the laboratory and constraints 4.17 the ones that do not.

$$u_{ikr} - u_{jkr} + \left(b_i + \tau_i + t_{ij} - a_j\right) \cdot x_{ijkr} \leq b_i - a_j \qquad \begin{array}{l} r = 2, \dots, R;\ \ k = 1, \dots, K; \\ j = 1, \dots, p+1;\ \ i = 0, \dots, p \end{array} \qquad (4.16)$$

$$u_{ikr} - u_{jkr} + \left(b_i + \tau_i + t_{ij} - a_j\right) \cdot x_{ijkr} \leq b_i - a_j \qquad \begin{array}{l} r = 1;\ \ k = 1, \dots, K;\ \ j = 1, \dots, p+1 \\ i = 1, \dots, p, p+2, \dots, p+d+1 \end{array} \qquad (4.17)$$

9) Waiting time in the lab between routes

Route r of vehicle k must start at least later than the arrival of its route $k-1$ plus the waiting time in the laboratory between routes τ_0.

$$u_{0kr} \geq u_{p+1,k,r-1} + \tau_0 \qquad k = 1, \dots, K;\ \ r = 2, \dots, R \qquad (4.18)$$

10) Maximum duration of each route

Constraints 4.19 and 4.20 state that the time that takes to return to the depot after visiting a given node is restricted so that the samples' lifetime is satisfied. Constraints 4.19 refers to the routes that start in the laboratory and constraints 4.20 to the ones that do not:

$$u_{p+1,k,r} - u_{jkr} \leq T_{max}^{j} + T_k \cdot (1 - \sum_{i=0}^{p} x_{ijkr}) \qquad r = 2, \dots, R;\ \ k = 1, \dots, K;\ \ j = 1, \dots, p \qquad (4.19)$$

$$u_{p+1,k,r} - u_{jkr} \leq T_{max}^{j} + T_k \cdot (1 - \sum_{i=1\&\ i\neq p+1}^{p+d+1} x_{ijkr}) \qquad r = 1;\ \ k = 1, \dots, K;\ \ j = 1, \dots, p \qquad (4.20)$$

11) Ensures that each vehicle has a maximum time limitation on the working day

Since each vehicle has a limitation on the length on the working day, constraints 4.21 set the maximum duration of vehicle k which has to be less than T_k.

$$u_{p+1,kr} - u_{ik1} \leq T_k \qquad r > 1;\ \ k = 1, \dots, K;\ \ i = p+2, \dots, p+d+1 \qquad (4.21)$$

12) Ensures that each collection request arrives in the laboratory in a given time period

The following constraints, constraints 4.22, state that the collection request i will arrive in the laboratory within a given time period t:

$$\sum_{k=1}^{K} \sum_{r=1}^{R} \sum_{t=0}^{T} y_{itkr} = 1 \qquad i = 1, \dots, p \qquad (4.22)$$

13) Decision variable *y* restrictions

Constraints 4.23 ensure that if collection request i has been visited by the r^{th} route of vehicle k the variable y_{itkr} can take the value of 1:

$$y_{itkr} \leq \sum_{j=1}^{p+1} x_{ijkr} \qquad \begin{array}{l} i = 1, \dots, p;\ \ t = 0, \dots, T; \\ k = 1, \dots, K;\ \ r = 1, \dots., R \end{array} \qquad (4.23)$$

14) Force relation between flow and y variable and discretize time into periods of units of time

Constraints 4.24 and 4.25 strengthen the connection within flow and y variables and divide time into periods of ω units of time. When using these constraints, in the case where $y_{itkr} = 1$ we have $\omega \cdot t \leq u_{p+1kr} \leq \omega \cdot (t+1)$, meaning that the laboratory has to be visited within the t^{th} time period, therefore within time $\omega \cdot t$ and $\omega \cdot (t+1)$.

$$u_{p+1,kr} < \omega \cdot (t+1) + M \cdot (1 - y_{itkr}) \qquad i = 1, \dots, p; \;\; t = 0, \dots, T; \;\; k = 1, \dots, K; \;\; r = 1, \dots, R \tag{4.24}$$

$$u_{p+1,kr} \geq \omega \cdot t - M \cdot (1 - y_{itkr}) \qquad i = 1, \dots, p; \;\; t = 0, \dots, T; \;\; k = 1, \dots, K; \;\; r = 1, \dots, R \tag{4.25}$$

15) To calculate the maximum number of boxes arriving in the laboratory during the available time periods

Constraints 4.26 calculate the laboratory maximum workload during the available time periods. In fact, and as expressed by constraints 4.26 the number of sample boxes that arrive in the laboratory in a specific time period is the same as the number of SCCs visited by the routes that returned to the laboratory in that same period. Supposing that a box of samples is gathered at each SCC request the workload is the number of boxes that arrive at the laboratory within this time period.

$$w \geq \sum_{k=1}^{K} \sum_{r=1}^{R} \sum_{i=1}^{p} y_{itkr} \qquad t = 0, \dots, T \tag{4.26}$$

16) To ensure that the objective function timings are correct

Since the objective function considers the starting time of the route of each vehicle the time of the laboratory, one has to guarantee that for the first route of each vehicle (the ones that may not start in the laboratory) that this value is equal to the time they depart from that first location, therefore constraints 4.27 are implemented in the following way:

$$u_{0kr} = u_{ikr} \qquad r = 1; \;\; i = p+2, \dots, p+d+1; \;\; k = 1, \dots, K \tag{4.27}$$

Moreover, constraints 4.28 enforce that the arrival time in the laboratory of a given route r occurs after its departing time:

$$u_{0kr} \leq u_{p+1kr} \qquad r = 2, \dots, R; \;\; k = 1, \dots, K \tag{4.28}$$

17) Lower bound on the minimum value of variable u

Constraints 4.29 define a lower bound on the minimum value of variable u. In fact, when visiting the arc (i,j) by the r^{th} route of vehicle k the variable x_{ijkr} equals 1 and consequently, $u_{ikr} + \tau_i + t_{ij} \leq u_{jkr}$. In case that x_{ijkr} equals 0 we have $u_{jkr} \geq a_j$ which is always valid.

$$u_{ikr} - \sum_{j=0 \& i \neq j \& j \neq p+1}^{p+d+1} (a_j + \tau_j - a_i + t_{ji}) \cdot x_{jikr} \geq a_i \qquad i = 1, \dots, p+1; \;\; r = 1, \dots, R; \;\; k = 1, \dots, K \tag{4.29}$$

18) Symmetry breaking constraints

Constraints 4.30 and 4.31 are based on the symmetry breaking constraints developed by Coelho and Laporte in 2014 [56]. Essentially, by ordering the use of the vehicles and the assignment of request to vehicles these constraints can remove symmetric solutions.

$$\sum_{j=1}^{p}\sum_{i=p+2}^{p+d+1} x_{ijk1} - \sum_{j=1}^{p}\sum_{i=p+2}^{p+d+1} x_{ijk-1,1} \leq 0 \qquad k = 2, \dots, K \qquad (4.30)$$

$$\sum_{r=1}^{R}\sum_{i=1}^{p} x_{ijkr} - \sum_{l=1}^{j-1}\sum_{r=1}^{R}\sum_{i=1}^{p} x_{ilk-1r} \leq 0 \qquad j = 1, \dots, p;\ \ k = 2, \dots, K \qquad (4.31)$$

19) Symmetry defeating constraints

Constraints 4.32 and 4.33 are symmetry defeating constraints which break the symmetry caused by variables y enhancing the model. Constraints 4.32 state that if the first route of vehicle $k-1$ has not been used, then the requests cannot be carried out by routes of the vehicle k. In addition, 4.33 states that route r of vehicle k can only be used if to route $r-1$ of this same vehicle is already associated at least one request.

$$\sum_{i=1}^{p}\sum_{r=1}^{R}\sum_{t=0}^{T} y_{itkr} \leq M \cdot \sum_{i=1}^{p}\sum_{t=1}^{T} y_{itk-1,1} \qquad k = 2, \dots, K \qquad (4.32)$$

$$\sum_{i=1}^{p}\sum_{t=0}^{T} y_{itkr} \leq M \cdot \sum_{i=1}^{p}\sum_{t=1}^{T} y_{itk,r-1} \qquad k = 1, \dots, K;\ \ r = 2, \dots, R \qquad (4.33)$$

21) Specific conditions

Constraint sets 4.34 and 4.35 state that one cannot travel from request i to request j when $a_i + \tau_i + t_{ij} > b_j$ and $a_i + \tau_j + t_{j,p+1} - b_j > T^i_{max}$. In fact, these equations eliminate redundant travels that violate the time windows and the maximum samples travel time.

$$\sum_{k=1}^{K}\sum_{r=1}^{R} x_{ijkr} = 0 \qquad \forall\, i,j \in V \setminus \{0, p+1\} \mid (a_i + \tau_i + t_{ij} > b_j) \qquad (4.34)$$

$$\sum_{k=1}^{K}\sum_{r=1}^{R} x_{ijkr} = 0 \qquad \forall i,j \in V \setminus \{0, p+1\} \mid (a_i - b_j + \tau_j + t_{j,p+1} > T^i_{max}) \qquad (4.35)$$

22) Relation between earliest visit time and possible time periods

Constraints 4.36 exhibits the connection concerning the earliest visit time of a given request and the possible time periods where the corresponding samples can arrive in the laboratory. Essentially, when $a_i + \tau_i + t_{i,p+1} \geq \omega t$, request i cannot arrive to the laboratory earlier than the t^{th} time period.

$$\sum_{k=1}^{K}\sum_{r=1}^{R}\sum_{l=1}^{t-1} y_{ilkr} = 0 \qquad \begin{array}{c} \forall\, i \in V\ \ \{0, p+1\} \mid \left(a_i + \tau_i + t_{i,p+1} \geq \omega \cdot t\right); \\ t = 1, \dots, T-1 \end{array} \qquad (4.36)$$

In summary, the BSTP model presented is composed by the objective function (4.5) and by all the restrictions from (4.6) to (4.36). The goal is to find the sequency and the visiting times of the SCCs that minimize both the maximum number of boxes arriving in the busiest period and the total travelling time while taking into consideration the SCCs time-windows and the maximum travel time for each route and each vehicle.

4.3 Chapter conclusions

In this chapter, the problem found in the case study is restated, and a mathematical model for the BSTP with desynchronized arrivals and open routes is developed. This model aims at designing new routes taking into consideration both the minimization of the routes total travel time as well as desynchronizing the arrivals in the laboratory. The model also considers several constraints that capture restrictions related to the desired plan (e.g., maximum time limitation for each vehicle and maximum time limitation for each route).

This model is in fact a Mixed Integer Linear Program (MILP) model and it was then implemented in the mathematical modelling software GAMS. Nevertheless, as it is presented in Chapter 6, the solver was not able to tackle the larger, real-life inspired instances efficiently, which motivated the development of an approximated method. In the next chapter, Chapter 5 the approach to solve this problem is discussed in detail.

5 Heuristic algorithm

As stated in the literature review the type of problems as the one under study hold various characteristics that make it difficult to be solved through the use of exact solution approaches, as the one used by GAMS software. This was observed when testing the model with several instances, as it is shown in Chapter 6. From this, arises the need of developing a heuristic algorithm that better addresses the problem, without compromising the results quality. Thus, a multi-start algorithm was developed. This used as basis the work of Naji-Azimi *et al.* (2016), which was adapted to fit the characteristics found in the case study chapter.

The metaheuristic used for the purpose of this thesis is explained in detail throughout the present chapter. In section 5.1, the algorithm and its components are detailed, and section 5.2 describes the way the visiting times are updated and how the time windows and maximal sample transportation time constraints are respected.

5.1 Multi-start heuristic algorithm

Multi-start algorithms hold two phases: the first one where an initial solution is generated and a second one where the generated solution is commonly, but not always, improved [57]. In the specific case of this algorithm, the first phase comprises the *Construction* procedure while the second phase is constituted by the *Extraction-Reinsertion* and *Swap* procedures. A schematic representation of the developed algorithm is represented in the following figure, Figure 5.1.

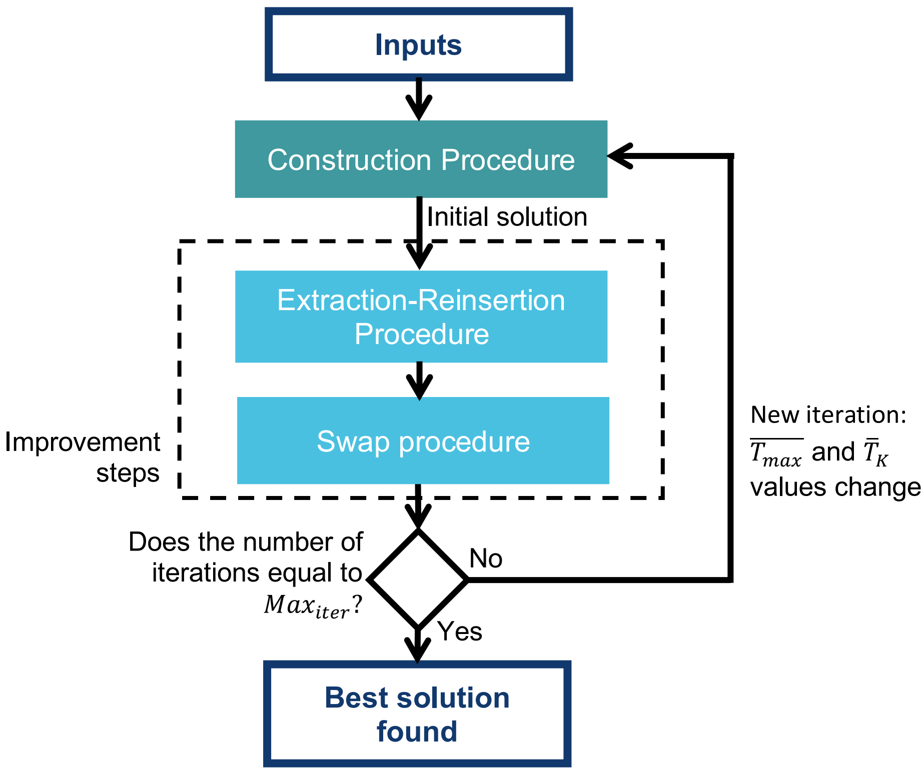


Figure 5.1 – Schematic representation of the multi-start algorithm developed.

As one can verify from Figure 5.1, the process repeats itself depending on the number of iterations chosen, Max_{iter}. Thus, to guarantee that diverse iterations lead to different solutions, a level of randomization is applied in two parameters: in the maximal sample transportation time (T_{max}) and in the maximum length of a vehicle working day (T_K). Both parameters influence the *Construction* procedure, which creates different feasible or unfeasible solutions. Moreover, these two parameters are adapted in conformity with the feasibility of the solution and only solutions that are feasible can pass to the improvement steps. Algorithm 1 provides an insight on the general steps of the multi-start algorithm used.

Algorithm 1: Multi-start algorithm

$BestSolution = \emptyset$
$Temp_{T_k} = T_k$
$Temp_{T_{max}} = T_{max}$
For $Iter = 1\ to\ Max_{iter}$ **do**
  Repeat
    $\overline{T_k} = random\left(0,\ T_k - Temp_{T_k}\right) + Temp_{T_k}$
    $\overline{T_{max}} = random\ (0, T_{max} - Temp_{T_{max}}) + Temp_{T_{max}}$
    $CurrentSolution =$ **Construction (**$\overline{T_k},\ \overline{T_{max}}$**)**
    If $CurrentSolution$ is feasible **then**
      $Temp_{T_{max}} = Temp_{T_{max}} - \alpha \cdot T_{max}$
      $Temp_{T_k} = Temp_{T_k} - \alpha \cdot T_k$
    Else
      $Temp_{T_{max}} = Min\ \{Temp_{T_k} + \alpha \cdot T_k;\ \ T_k\}$
      $Temp_{T_k} = Min\ \{Temp_{T_{max}} + \alpha \cdot T_{max};\ \ T_{max}\}$
    End If
  Until $CurrentSolution$ is feasible
  $CurrentSolution =$ **Extraction-Reinsertion (**$CurrentSolution$**)**
  $CurrentSolution =$ **Swap (**$CurrentSolution$**)**
  If $CurrentSolution$ improves the cost of the best-known solution, **then**
    $BestSolution = CurrentSolution$
End for

From Algorithm 1 it is possible to verify that this multi-start algorithm consists in two loops: the outer loop (the *For* loop) and the inner loop (the *Repeat – Until* loop). The inner loop consists in constructing an initial feasible solution; thus, the *Construction* procedure is applied with distinct temporary values for the maximum vehicle travel time ($\overline{T_k}$) and the sample travel time ($\overline{T_{max}}$) parameters until a feasible solution is found. The outer loop comprises the inner loop and the two improvement steps. In the first iteration, $Temp_{T_k}$ and $Temp_{T_{max}}$ are initialized by setting their values to T_k and to T_{max}, respectively, which have to be the first initial feasible parameters. After the first iteration, if the solution obtained by the *Construction* procedure is feasible, the both $Temp_{T_k}$ and $Temp_{T_{max}}$ parameters reduce by a factor α which is another input variable (the value of $\alpha = 0.1$ was used for the purpose of this thesis):

$$Temp_{T_k} = Temp_{T_k} - \alpha \cdot T_k \tag{5.1}$$

$$Temp_{T_{max}} = Temp_{T_{max}} - \alpha \cdot T_{max} \tag{5.2}$$

On the other hand, if the solution is not feasible in order to increase the chance of finding a feasible solution the values of these two parameters are increased by applying:

$$Temp_{T_k} = Min\,\{Temp_{T_k} + \alpha \cdot T_k;\ \ T_k\} \tag{5.3}$$

$$Temp_{T_{max}} = Min\,\{Temp_{T_{max}} + \alpha \cdot T_{max};\ \ T_{max}\} \tag{5.4}$$

Finally, the temporary values used in the *Construction* procedure to be used as the maximum travel time and maximum sample time are:

$$\overline{T_k} = random\,\left(0,\, T_k - Temp_{T_k}\right)\, +\, Temp_{T_k} \tag{5.5}$$

$$\overline{T_{max}} = random\,\left(0,\, T_{max} - Temp_{T_{max}}\right)\, +\, Temp_{T_{max}} \tag{5.6}$$

, where $random(0,\ x)$ is an integer number randomly selected between 0 and x.

After obtaining an initial feasible solution the improvement procedures are now applied. Thus, as previously mentioned, the *Extraction-Reinsertion* and the *Swap* procedures attempt to improve the quality of the solution coming from the *Construction* procedure. To do so, the original values of the sample travel time and the vehicle travel time are used. After the maximum number of iterations has been achieved the algorithm stops and the best solution found is returned.

5.1.1 Phase one – Generation of an initial solution

Construction procedure

As stated, the construction procedure forms an initial solution, where the nodes (i.e., the SCCs) are sequentially added to the routes. Routes are initialized by adding their respective starting and finishing points.

To find the method to select the first node to be visited by a route, n_1, tests were performed on 5 instances whose characteristics are depicted in Table 5.1. These tests lied on finding which approach would give the best results. More specifically, the following two approaches were tested:

- **Approach 1**: n_1 is chosen randomly from the set of request nodes whose travel time is the greatest to the laboratory and whose upper bond time-window is the lowest;
- **Approach 2:** n_1 is chosen randomly from the set of request nodes who are the closest to the departing point and whose upper bond time window is the lowest.

Table 5.1 – Tests to find the methodology to choose the first node to be visited by a route.

#	Instances characteristics							Objective function results	
	General			SCCs					
	V'	P	T_{max}	P/V'	$\bar{t}_{ij}$ (min)	Std. a_j	Std. b_j	Approach 1	Approach 2
1	5	5	300	1	15.19	13.42	53.67	243	243
2	12	12	300	1	23.49	23.79	0	541	545
3	18	18	300	1	9.74	156.82	148.01	552	553
4	21	29	300	1.4	5.87	147.33	173.68	842	847
5	39	50	300	1.28	10.02	129.09	141.11	1254	1380

As one can verify from the results obtained in Table 5.1, the method that leads to better results (i.e., that minimizes the most the objective function) is in fact Approach 1 which is the one that selects the node that is farthest from the laboratory with the lowest upper bond. This method provided for all instances an equal or better performance that the second approach, Approach 2. Thus, to select the first node to be visited, n_1, the following set of rules are used:

- $N_1 = argmax_{i \in P}\{t_{i,p+1}\}$, i.e., the set of request nodes $i \in P = \{v_1, v_2, \dots, v_p\}$ whose travel time is the greatest to the laboratory;
- $N_2 = argmin_{i \in N_1}\{b_i\}$, i.e., the set of request nodes $i \in N_1$ whose time window upper bond is the lowest;
- A random node, n_1, is selected from N_2.

After inserting the first node, n_1, in the working route, to insert a new node certain rules are to be followed. Firstly, a node from the unvisited set of nodes is chosen in sequence and the possibility of adding it in every position of that route is evaluated. If there are feasible placing positions, the node will be placed in the position that results in the smallest increase in the route's total travel time. For example, if node i is placed between nodes j and k the detour is calculated as $t_{ji} + t_{ik} - t_{jk}$. In case there is not a feasible position for that node in that route, the algorithm tries to place the next unvisited node in that same route. The process repeats until all unvisited nodes are tested to be placed in that route. Afterwards, the same process occurs for the other unfilled routes. The described procedure is now summarized in Algorithm 2.

Algorithm 2: Construction procedure

For all $k \in K$ **do**

For all $r \in R$ **do**

Begin by initiate route r of vehicle k with the first node, n_1

While there is a node not yet placed to be added in a feasible position **do**

Add to the route r, if feasible, the nodes in the positions with the smallest increase in the route's total duration

End While

$r = r + 1$

End for

$k = k + 1$

End for

5.1.2 Phase two – Improvement of the initial solution

Extraction-Reinsertion procedure

The first improvement procedure, the *Extraction-Reinsertion* procedure, has the purpose of reducing the value of the objective function by repositioning some of the nodes of the solution previously obtained in the *Construction* procedure. Hence, to calculate this cost, the length of the time period, ω, has now to be considered.

The procedure lies on taking every node and placing it on its best feasible position. Thus, starting from first route of the first vehicle all the nodes are repositioned in all possible locations. After trying to place a node in all positions the node is placed in the feasible position that leads to the best improvement, i.e., it is placed it in the position that minimizes the most the cost of the objective function. This procedure is repeated until no further improvement can be reached. During the procedure, a new route can be created, or an already existing route can be closed.

Swap procedure

The second improvement procedure, the *Swap* procedure, also aims at minimizing the cost of the current solution, i.e., the one coming from the *Extraction-Reinsertion* procedure. To improve the solution, following the order of the nodes in the current solution, each pair of nodes is considered, and their corresponding positions are swapped. The swap is applied to all possible combination of two nodes over all the vehicles' routes. As soon as a move improves the solution's cost, it is accepted. This procedure stops whenever the swapping of all the available nodes offers no more improvement.

5.2 Information update

In all the three procedures previously described, information regarding the time windows and the maximum sample transportation time when adding or removing a node from a given position need to be updated so that only feasible solutions are generated. Therefore, in order to efficiently update the time windows and the maximum sample transportation time constraints, rules to update the earliest and latest visiting time of the nodes in the solution need to be taken into consideration. Thus, an adaptation of the method of Campbell and Savelsbergh (2004) for the VRP with time windows was developed [58].

The r^{th} route of vehicle k is represented by $k_r = \{c_0^{k_r}, c_1^{k_r}, \ldots, c_i^{k_r}, c_{i+1}^{k_r}, \ldots, c_{n_{k_r}}^{k_r}, c_{n_{k_r}+1}^{k_r}\}$ where n_{k_r} is the number of nodes visited by the r^{th} route of vehicle k, $c_0^{k_r}$ is the departing location node of vehicle k and route r and $c_{n_{k_r}+1}^{k_r}$ the lab, the arrival node. The earliest and latest time at which node $c_j^{k_r}$ can be visited by the route r^{th} route of vehicle k is represented by $E_{c_j^{k_r}}$ and $L_{c_j^{k_r}}$, respectively.

Initially, for each k and r the earliest and latest time of both the departing and arrival nodes, $E_{c_0^{k_r}}$ and $E_{c_{n_{k_r}+1}^{k_r}}$, and, $L_{c_0^{k_r}}$ and $L_{c_{n_{k_r}+1}^{k_r}}$ respectively, is set to be the same as their corresponding upper and lower window. In case of inserting a node $c_j^{k_r}$ between nodes $c_i^{k_r}$ and $c_{i+1}^{k_r}$ the following two relations are applied:

$$E_{c_j^{k_r}} = Max\,\{a_{c_j}^{k_r},\, E_{c_i^{k_r}} + \tau_{c_i}^{k_r} + t_{c_i^{k_r} c_j^{k_r}}\} \tag{5.7}$$

$$L_{c_j^{k_r}} = Min\,\{b_{c_j}^{k_r},\, L_{c_{i+1}^{k_r}} - \tau_{c_j}^{k_r} - t_{c_j^{k_r} c_{i+1}^{k_r}}\} \tag{5.8}$$

If $E_{c_j^{k_r}} \leq L_{c_j^{k_r}}$, introducing the node $c_j^{k_r}$ between nodes $c_i^{k_r}$ and $c_{i+1}^{k_r}$ will not violate the time window constraints. Nevertheless, it is still mandatory to verify both the maximum duration of the vehicle length and the maximum sample travel time are satisfied. Therefore, the following relations, equation 5.9 and 5.10, dictate how to update the earliest and latest start time of nodes after and before $c_j^{k_r}$:

$$E_{c_s^{k_r}} = Max\,\{a_{c_s}^{k_r},\, E_{c_{s-1}^{k_r}} + \tau_{c_{s-1}}^{k_r} + t_{c_{s-1}^{k_r} c_s^{k_r}}\} \tag{5.9}$$

$$L_{c_s^{k_r}} = Min\,\{b_{c_s}^{k_r},\, L_{c_{s+1}^{k_r}} - \tau_{c_s}^{k_r} - t_{c_s^{k_r} c_{s+1}^{k_r}}\} \tag{5.10}$$

Moreover, the minimum length of the r^{th} route of vehicle k requires serving the route at the latest feasible time, by setting it to $vt_{c_{n_{k_r+1}}}^{k} - vt_{c_0}^{k_r}$. In this expression $vt_{c_i}^{k_r}$ is the real visit time of node $c_i^{k_r}$ by the route r of vehicle k and can be calculated as:

$$vt_{c_i}^{k_r} = Max\,\{a_{c_i}^{k_r}, vt_{c_{i-1}}^{k_r} + \tau_{c_{i-1}}^{k_r} + t_{c_{i-1}^{k_r}, c_i^{k_r}}\} \qquad i = 1, \ldots, n_{k_r} + 1 \tag{5.11}$$

In addition, $vt_{c_0}^{k_r} = L_{c_0^{k_r}}$.

Equations 5.12 and 5.13 ensure that the maximum vehicle travel time and the maximum sample travel time are satisfied:

$$vt_{c_{n_{k_R}+1}}^{k_R} - vt_{c_o}^{k_1} \leq T_k \tag{5.12}$$

$$vt_{c_{n_{k_r}+1}}^{kr} - vt_{c_1}^{k_r} \leq T_{max} \qquad r = 1, \dots, R \tag{5.13}$$

Finally, $vt_{c_0^{k_r}} = vt_{c_{n_{k_{r-1}}+1}^{k_{r-1}}} + \tau_0$ for $r = 2, \dots, R$.

5.3 Chapter conclusions

The present chapter describes in detail the proposed metaheuristic algorithm to solve complex real-life problems of the BSTP with desynchronized arrivals and open routes. Particularly, this algorithm comprises two main phases: the creation of an initial solution and the improvement of this solution. Several different initial solutions are created based on the number of iterations chosen for the algorithm. Randomization in the creation of initial solution is ensured by applying different initial values at each iteration for the maximal sample transportation time (T_{max}) and the maximum length of a vehicle working day (T_K).

The objective is to apply this algorithm to the case study problem previously presented in Chapter 2. However, one still has to ensure that the multi-start algorithm provides quality results and to find the parameters to apply to the case study have to be established (namely, the penalization to consider, θ, the way to discretize time periods, ω, and the number of iterations). Hence, the next chapter, Chapter 6, provides results for both the mathematical model using the optimization program GAMS and the heuristic algorithm in a set of created instances.

6 Performance numerical results

The purpose of this chapter is to present the numerical results obtained by applying the solution approaches in a set of created instances to analyse their performance. Hence, both the mathematical model using an optimization program and the heuristic method are tested and compared on different examples.

The chapter is organized as follows: section 6.1 describes the set of instances created based on the entity's data; section 6.2 reports the results obtained by running the linear model on GAMS software with a time limitation of 3 hours. Section 6.3 presents the results obtained by the heuristic method, the multi-start algorithm, and compares them with the ones obtained in section 6.2. Finally, section 6.4 presents a conclusion for this chapter.

6.1 Dataset presentation

As previously mentioned, the entity under study has pre-defined routes that are programmed to occur in every day of the week, and which are organized based on the Specimen Collection Centres (SCCs) geographical locations. Each one of these SCCs can be visited more than once, having one or more collection requests per day. In order to test the mathematical model and the algorithm developed, a set of 15 instances was designed based on the geographical locations considered by the entity in region of A to construct their routes. These 7 regions are: Region A1, Region A2, Region A3, Region A4, Region A5, Region A6 and Region A7.

The created instances were classified into three groups based on their size: small (less than 20 collection requests), medium (more or equal than 20 but less than 50 collection requests) and large (more or equal than 50 collection requests). Table 6.1 summarizes the basic characteristics of these instances, namely the minimum and maximum number of SCCs and the minimum and maximum number of collection requests depending on the instances' size.

Table 6.1 – Instances summary.

Size	Number of instances	Number of SCCs		Number of collection requests	
		Min.	Max.	Min.	Max.
Small	5	5	15	5	15
Medium	5	18	25	18	29
Large	5	39	70	50	71

In order to define the instances, information regarding the SCCs was provided, namely their names and addresses, the time-windows at which they should be visited and the average maximal transportation time for the samples. Information regarding the lunch hours and the maximum time of a working day was also provided as well as the departing places of each vehicle.

With the information regarding the SCCs, in particular their names and addresses, it was possible to obtain the time matrices necessary to comprehend the spatial density related to the considered instances. Firstly, with the tool “Google My Maps” the geolocations (i.e., latitude and longitude) of all the significant points were obtained. Afterwards, with the geolocations defined it was possible to obtain the distance matrices with the use of a Google Maps API [59]. Then, it was asked to an expert what the value of the average velocity should be considered to map the time matrix. It was responded that for Region A1, Region A2, Region A3, Region A4 and Region A5 the average velocity should be around 60 km/h while for Region A6 and Region A7 the average velocity should be about 75 km/h.

Moreover, to dictate the value of the routes’ maximal transportation time, the same worker stated that it should be around 5 hours. Except for the regions where the SCCs are more far away from one another, which can take more than 5 hours. Thus, for most places T_{max} was considered to be 5 hours (i.e., 300 minutes) with the exception of Region A2.2, Region A6 and Region A7 which can take higher values than 5 hours (up to 540 minutes).

Another information that had to be provided was the desired time windows to pass in each SCC. These time windows were defined in a way that guarantees that all the biomedical samples arrive to the laboratory in that day.

Table 6.2 reports a summary of the properties of the 15 instances created. It is important to highlight that the columns headings follow the same notation as the one used in Chapter 4. The names of the instances are based on the previously presented geographical regions. The numbers after the regions’ name represent different sets of SCCs of the same geographical region. Moreover, it specifies for each instance the number of SCCs (V'), the number of collection requests (P) and the maximum considered transportation time in minutes (T_{max}). To characterize the SCCs of each instance the table also provides information on: the average number of pickups for the SCCs (P/V'), the average travel time in minutes ($\bar{t}_{ij}$) and the standard deviations on the opening and closing hours in minutes (respectively, $std.a_j$ and $std.b_j$).

Table 6.2 – Summary of the properties of the 15 instances.

#	Size	Name	General			SCCs			
			V'	P	T_{max}	P/V'	$\bar{t}_{ij}$ (min)	Std. a_j	Std. b_j
1	Small ($P < 20$)	Region A3.1	5	5	300	1	15.19	13.42	53.67
2		Region A5.2	10	10	300	1	11.15	34.79	0
3		Region A2.3	12	12	300	1	23.49	23.79	0
4		Region A4.1	15	15	300	1	6.18	50.03	0
5		Region A4.2	15	15	300	1	6.77	48.73	0
6	Medium ($20 \leq P < 50$)	Region A2.1	21	29	300	1.4	5.87	147.33	173.68
7		Region A2.2	21	21	350	1	25.00	151.96	107.17
8		Region A1.3	18	18	300	1	9.74	156.82	148.01
9		Region A5.1	21	21	300	1	16.88	144.90	115.54
10		Region A3.2	25	26	300	1.04	13.85	149.94	141.29
11	Large ($P \geq 50$)	Region A2.1 + A2.2	42	50	350	1.19	18.10	149.10	148.08
12		Region A2.1 + A2.4	39	50	300	1.28	10.02	129.09	141.11
13		Region A3.1 + A3.2 & Region A5.1 + A5.2	70	71	300	1.01	15.62	127.26	93.26
14		Region A6.1 + A6.2 & Region A7.1 + A7.2 + A7.3	62	62	450	1	53.43	133.61	102.92
15		Region A1.1 + A1.4	41	51	300	1.24	6.00	136.06	110.8

Finally, to fully characterize the 15 instances there are general parameters that can be applied to all of them. These parameters are:

- **Loading time in the SCCs (τ_i) and unloading time in the lab (τ_0)**:
  Experts referred that the loading time in the SCCs should be around 10 minutes. The unloading time at the laboratory, on the other hand, should set to be about 60 minutes since it is in this period that the drivers have their lunch break.

- **Working hours in the laboratory (a_0 and b_0)**:
  Entity experts also stated that the samples should arrive in the lab between 8h to 21h.

- **Large constant (M):**
  In the formulation of the model a large constant, M, is used and it is fixed to be 100000. It is important to note that this value is larger than any other value used such as distances or times.

A summary of these common parameters can be found in the following table, Table 6.3.

Table 6.3 – Common parameters.

τ_0 (min)	τ_i (min)	a_0 (min)	b_0 (min)	M
60	10	480	1260	100000

6.2 Optimization solver results

As referred in Chapter 4, the developed model was tested using the optimization software GAMS. The properties of the computer used for this purpose are stated in the following table, Table 6.4.

Table 6.4 – System properties

System properties	
2 Processors	Intel(R) Xenon(R) CPU & X5680 @ 3.33GHz
Installed Memory (RAM)	24 GB

Several tests were performed on the 15 instances previously described to test the mathematical model and to find the parameters that would fit better the case study context. For all these tests, the computational time was limited to 3 hours (10800 seconds). Additionally, for all the tests that the optimization program could not find any solution within the time limitation stipulated (3 hours of computation) were marked with an asterisk.

As previously mentioned in Chapter 4, the aim of the optimization program is to find a solution that minimizes the objective function. This function represents the sum of the routes' total duration with a penalization, θ, associated with the number of boxes arriving in the laboratory in the most occupied period, which is represented by w. To measure the number of boxes arriving in the busiest period, the day is discretized into intervals of ω minutes. The objective is to find which values of θ and ω fit best the entity under study characteristics, thus, two different experiments are be performed. The first lies on finding which penalization fits best the case study context: if $\theta = 25$, $\theta = 50$ or $\theta = 75$. The second one is to find the best form to discretize time: if in intervals of 30 minutes or 60 minutes (i.e., $\omega = 30$ or $\omega = 60$).

6.2.1 Penalization factors experiment (θ)

As mentioned, the first experiment lies on comparing the results for different penalties values to see which value is the one that better balances the minimization of the route's duration and the number of arrivals during the busiest period. For this purpose, it was considered that the time discretization should be $\omega = 60$. Thus, Table 6.5 reports the results for different values of penalties: $\theta = 25$, $\theta = 50$ and $\theta = 75$ and presents the values obtained for the objective function, routes' total duration (in minutes), number of boxes arriving in the busiest period (w), lower bound, gap and the computational time (in seconds).

Table 6.5 – Optimization Solver Results for the 15 instances with different values of θ ($\theta = 25$, $\theta = 50$ and $\theta = 75$) considering $\omega = 60$.

	$\theta = 25$						$\theta = 50$						$\theta = 75$					
	Objective Function	**Total time (min)**	w	**Lower Bound**	**Gap (%)**	**Computational Time (sec)**	**Objective Function**	**Total time (min)**	w	**Lower Bound**	**Gap (%)**	**Computational Time (sec)**	**Objective Function**	**Total time (min)**	w	**Lower Bound**	**Gap (%)**	**Computational Time (sec)**
1	241	116	5	241	0	0.03	366	116	5	366	0	0.09	491	116	5	491	0	0.11
2	413	163	10	413	0	2787	663	163	10	663	0	2609	913	163	10	913	0	2801
3	541	241	12	300	44.5	10800	841	241	12	600	28.7	10800	1141	241	12	900	21.1	10800
4	566	191	15	375	33.7	10800	941	191	15	750	20.3	10800	1316	191	15	1125	14.5	10800
5	577	202	15	375	35	10800	952	202	15	750	21.2	10800	1328	203	15	1125	15.3	10800
6	840	415	17	54	93.6	10800	1264	414	17	120	90.2	10800	*	*	*	*	*	10800
7	*	*	*	*	*	10800	*	*	*	*	*	10800	*	*	*	*	*	10800
8	540	290	10	100	81.5	10800	790	290	10	350	55.7	10800	1040	290	10	600	42.3	10800
9	698	323	15	50	92.8	10800	1061	361	14	100	90.6	10800	1202	377	11	255	81.2	10800
10	735	360	15	54	92.6	10800	1114	364	15	100	91	10800	*	*	*	*	*	10800
11	*	*	*	*	*	10800	*	*	*	*	*	10800	*	*	*	*	*	10800
12	*	*	*	*	*	10800	*	*	*	*	*	10800	*	*	*	*	*	10800
13	*	*	*	*	*	10800	*	*	*	*	*	10800	*	*	*	*	*	10800
14	*	*	*	*	*	10800	*	*	*	*	*	10800	*	*	*	*	*	10800
15	*	*	*	*	*	10800	*	*	*	*	*	10800	*	*	*	*	*	10800

Analysing Table 6.5, one can verify that once the instances grow the optimization program stops finding the optimal solution. The program only found the optimal result for the first two small instances. Further, as the complexity of the problem increases the optimization program ceases finding results within the 3 hours of computation. This occurred for all the large size instances and for instance 7. It is also important to point out that for the penalty factor of $\theta = 75$ the model could not find a solution for instances 6 and 10, while it found for the other two penalty factors.

Comparing the different results for different values of penalty, it is possible to perceive that the program only found reasonable values of gap for the small instances, while for the medium size instances the values of gap increased in a significant amount. In addition, analysing each individual instance, for the same computational time the values of gap decrease when the penalty increases. This was in fact verified for all the instances where GAMS could find results.

By evaluating the routes' total time for when the optimization program was able to dictate results, it is possible to verify that the number of boxes arriving in the laboratory in the busiest period does not change drastically within the penalties. The most pertinent instance to analyse is instance 9, since it provided different values of w for all the different penalties. For the penalty of $\theta = 25$, it presented that the routes' total time would be 323 minutes with 15 boxes arriving in the busiest period, for the penalty of $\theta = 50$ one would have a value of the routes' total time of 361 minutes with 14 boxes arriving in the lab in the busiest period and, finally, for the penalty of $\theta = 75$, it gave that the routes' total time would be 377 minutes with 11 boxes arriving in the most occupied time period. These results were, as expected, that higher values of penalty would provide higher values for the routes' total time and decrease the number of boxes arriving in the laboratory in the busiest period. The main reason why this reality did not occur in the other instances is due to the time windows restrictions of the collection requests. In fact, some collection requests time windows forced SCCs to be visited by a particular route not allowing the number of boxes arriving in the laboratory in the most occupied time period to change.

With all being said, the penalty that is found to fit better the purpose of this thesis is the one where $\theta = 50$, since it offers a favourable balance between the number of boxes arriving in the busiest period and the routes' total time. This value does not penalize to much the arrivals in the laboratory and is able to find favourable results for the routes' total time. Furthermore, it gives better gap results when comparing with the penalization of $\theta = 25$ and was able to provide results in some instances when the penalty of $\theta = 75$ was not able to provide. Thus, for the case study of Region A, a penalization of $\theta = 50$ is the value used to find new routes.

6.2.2 Time period discretization experiment (ω)

The second experiment lied on finding what would be the finest way to discretize time, if in intervals of 30 minutes or if intervals of 60 minutes. Therefore, Table 6.6 reports the numerical results produced for these instances when $\omega = 30$ and $\omega = 60$.

Table 6.6 – Optimization Solver Results for the 15 instances with different values of ω ($\omega = 30$ and $\omega = 60$) considering $\theta = 50$.

#	$\boldsymbol{\omega = 30}$						$\boldsymbol{\omega = 60}$					
	Objective Function	**Total time (min)**	$\boldsymbol{w}$	**Lower Bound**	**Gap (%)**	**Computational Time (sec)**	**Objective Function**	**Total time (min)**	$\boldsymbol{w}$	**Lower Bound**	**Gap (%)**	**Computational Time (sec)**
1	366	116	5	366	0	0.56	366	116	5	366	0	0.09
2	663	163	10	663	0	3214	663	163	10	663	0	2609
3	841	241	12	600	28.7	10800	841	241	12	600	28.7	10800
4	941	191	15	750	20.3	10800	941	191	15	750	20.3	10800
5	951	201	15	750	21.1	10800	952	202	15	750	21.2	10800
6	1266	416	17	50	96.1	10800	1264	414	17	120	90.2	10800
7	*	*	*	*	*	10800	*	*	*	*	*	10800
8	790	290	10	50	93.7	10800	790	290	10	350	55.7	10800
9	1062	362	14	50	95.3	10800	1061	361	14	100	90.6	10800
10	*	*	*	*	*	10800	1114	364	15	100	91	10800
11	*	*	*	*	*	10800	*	*	*	*	*	10800
12	*	*	*	*	*	10800	*	*	*	*	*	10800
13	*	*	*	*	*	10800	*	*	*	*	*	10800
14	*	*	*	*	*	10800	*	*	*	*	*	10800
15	*	*	*	*	*	10800	*	*	*	*	*	10800

Similar to the results previously presented in Table 6.5, also in Table 6.6 for both $\omega = 60$ and $\omega = 30$ the optimization program was not able to find results for the large instances and for instance 7 even after 3 hours of computational time. Furthermore, for instance 10 GAMS software could only find results when $\omega = 60$.

In terms of the objective function, the results provided by when the time is discretized into intervals of 30 minutes ($\omega = 30$) are similar to when the time is discretized into intervals of 60 minutes ($\omega = 60$). Nevertheless, when the optimization program did not find equal results, as one can notice in instances 6 and 9, the objective function values were higher for $\omega = 30$, while the number of boxes arriving in the busiest period were the same. Since the aim is to minimize the objective function, the quality of the results provided for $\omega = 30$ is worse when compared to $\omega = 60$.

Finally, with all things considered, it is possible to conclude that the complexity of solving the problem when the day is discretized into intervals of 30 minutes is much higher than when discretizing it into intervals of 60 minutes. Thus, since the performance of the ω parameter was remarkably higher for when it took the value of 60, for the purpose of this thesis and for the case study chapter, this parameter takes the value of 60.

In summary, using the optimization program to find solutions for the developed model is not viable for large instances, since it was not able to provide results even after several hours of computation. Nevertheless, from the results obtained, one can conclude that the parameters that fit best the purpose of the entity under study are when the penalty value is equal to 50 (i.e., $\theta = 50$) and when the time is discretized into intervals of 60 minutes (i.e., $\omega = 60$). As a result, these are the parameters to be used in Chapter 7 to compute Region A routes. For the case of Region B since the objective is not to desynchronize the arrivals a penalty of $\theta = 0$ is considered.

6.3 Heuristic results

From the previous section it was possible to perceive that the optimization program was not able to find results for the large size instances. Since the problem found in this entity has many SCCs, the need of developing a heuristic algorithm emerged; ergo, the multi-start heuristic algorithm described in Chapter 5 was developed and implemented in Python 3.8.1.

For both experiments presented in the previous section (in Table 6.5 and 6.6), even after multiple hours of computation GAMS software was not able to find reasonable values of lower bound to use to compare the objective function with. Thus, for the purpose of this section, the solutions obtained by GAMS are the ones used as benchmark and not the values of the lower bound. Thereby, the solutions obtained previously are used to compare to the ones obtained using the heuristic algorithm.

To test whether the developed heuristic is appropriate to be used for the purpose of the case study, a variety of tests are developed to compare the heuristic results with the ones obtained in the previous section. Also, a validation of the parameters chosen (more specifically, $\theta = 50$ and $\omega = 60$) is performed.

Thereafter, for this purpose, the same 15 instances are used. To be able to compare them, the results provided for the heuristic are executed using the same computer as the one used to obtain GAMS results, where the system properties are presented in Table 6.4.

The first experiment performed, presented in Table 6.7, depicts the results obtained for the optimization program, GAMS, and for the heuristic when different values of penalizations are considered, namely $\theta = 25$, $\theta = 50$ and $\theta = 75$. For this purpose, the parameter considered to discretize time is $\omega = 60$ and the number of iterations considered for the heuristic is 100. This table presents the objective function results, the routes' total time (in minutes), the number of boxes arriving in the busiest period (w), the difference in percentage (i.e., the deviation of the objective function of the heuristic compared to the GAMS solution) and the computational time (in seconds).

The second experiment, presented in Table 6.8, portraits the results of discretizing the day into different intervals, namely for when the day is discretized into intervals of 30 minutes ($\omega = 30$) and for when the day is discretized into intervals of 60 minutes ($\omega = 60$). This table introduces the results for both solution methodologies regarding: the value of the objective function, routes' total time (in minutes), number of boxes arriving in the busiest period (w), the difference in percentage (i.e., the deviation of the objective function of the heuristic compared to the GAMS solution) and computational time (in seconds). The number of iterations considered for the heuristic is also 100 and the penalization considered is $\theta = 50$.

Table 6.7 – Heuristic and optimization solver results for different values of θ ($\theta = 25$, $\theta = 50$ and $\theta = 75$) considering $\omega = 60$.

#	θ = 25									θ = 50									θ = 75								
	Objective function		Total time (min)		w		Difference (%)	Computational Time (sec)		Objective Function		Total time (min)		w		Difference (%)	Computational Time (sec)		Objective Function		Total time(min)		w		Difference (%)	Computational Time (sec)	
	G	H	G	H	G	H		G	H	G	H	G	H	G	H		G	H	G	H	G	H	G	H		G	H
1	241	243	116	118	5	5	0.8	0.03	38.8	366	368	116	118	5	5	0.5	0.09	39	491	493	116	118	5	5	0.4	0.11	39
2	413	420	163	170	10	10	1.7	2787	446	663	670	163	170	10	10	1.1	2609	449	913	920	163	170	10	10	0.8	2801	448
3	541	541	241	241	12	12	0	10800	439	841	841	241	241	12	12	0	10800	440	1141	1141	241	241	12	12	0	10800	442
4	566	567	191	192	15	15	0.2	10800	2418	941	942	191	192	15	15	0.1	10800	2440	1316	1317	191	192	15	15	0.08	10800	2445
5	577	580	202	205	15	15	0.5	10800	1399	952	955	202	205	15	15	0.3	10800	1392	1328	1330	203	205	15	15	0.2	10800	1383
6	840	842	415	417	17	17	0.2	10800	4387	1264	1267	414	417	17	17	0.2	10800	4410	*	1692	*	417	*	17	-	10800	4575
7	*	799	*	399	*	16	-	10800	1989	*	1199	*	399	*	16	-	10800	1991	*	1599	*	399	*	16	-	10800	2004
8	540	550	290	300	10	10	1.9	10800	1010	790	800	290	300	10	10	1.2	10800	1026	1040	1050	290	300	10	10	1	10800	1005
9	698	723	323	348	15	15	3.6	10800	1829	1061	1098	361	348	14	15	3.5	10800	1834	1202	1431	377	381	11	14	19.1	10800	2414
10	735	862	360	437	15	17	1.2	10800	2341	1114	1287	364	437	15	17	15.5	10800	2341	*	1712	*	437	*	17	-	10800	2351
11	*	1216	*	791	*	17	-	10800	15176	*	1603	*	753	*	17	-	10800	15936	*	2056	*	781	*	17	-	10800	16091
12	*	1198	*	773	*	17	-	10800	19829	*	1623	*	773	*	17	-	10800	19131	*	2048	*	773	*	17	-	10800	19874
13	*	1924	*	1149	*	31	-	10800	56862	*	2699	*	1149	*	31	-	10800	56126	*	3474	*	1149	*	31	-	10800	55717
14	*	2337	*	2037	*	12	-	10800	29122	*	2640	*	1990	*	13	-	10800	23148	*	2946	*	2046	*	12	-	10800	22722
15	*	1185	*	685	*	20	-	10800	14117	*	1685	*	685	*	20	-	10800	14147	*	2185	*	685	*	20	-	10800	14352

Table 6.8 – Heuristic and optimization solver results for different values of ω ($\omega = 30$ and $\omega = 60$) considering $\theta = 50$.

#	$\omega = 30$									$\omega = 60$								
	Objective Function		Total time (minutes)		w		Difference (%)	Computational Time (sec)		Objective Function		Total time (minutes)		w		Difference (%)	Computational Time (sec)	
	G	H	G	H	G	H		G	H	G	H	G	H	G	H		G	H
1	366	368	116	118	5	5	0.5	0.56	42	366	368	116	118	5	5	0.5	0.09	39
2	663	670	163	170	10	10	1.1	3214	497	663	670	163	170	10	10	1.1	2609	449
3	841	841	241	241	12	12	0	10800	486	841	841	241	241	12	12	0	10800	440
4	941	942	191	192	15	15	0.1	10800	2612	941	942	191	192	15	15	0.1	10800	2440
5	951	955	201	205	15	15	0.4	10800	1567	952	955	202	205	15	15	0.3	10800	1392
6	1266	1267	416	417	17	17	0.1	10800	5037	1264	1267	414	417	17	17	0.2	10800	4410
7	*	1199	*	399	*	16	-	10800	2177	*	1199	*	399	*	16	-	10800	1991
8	790	800	290	300	10	10	1.3	10800	1150	790	800	290	300	10	10	1.2	10800	1026
9	1062	1098	362	348	14	15	3.4	10800	1991	1061	1098	361	348	14	15	3.5	10800	1834
10	*	1287	*	364	*	17	-	10800	2603	1114	1287	364	437	15	17	15.5	10800	2341
11	*	1608	*	758	*	17	-	10800	18301	*	1603	*	753	*	17	-	10800	15936
12	*	1609	*	759	*	17	-	10800	22035	*	1623	*	773	*	17	-	10800	19131
13	*	2015	*	1215	*	16	-	10800	42942	*	2699	*	1149	*	31	-	10800	56126
14	*	2493	*	1893	*	12	-	10800	16707	*	2640	*	1990	*	13	-	10800	23148
15	*	1685	*	685	*	20	-	10800	14191	*	1685	*	685	*	20	-	10800	14147

Looking at the results provided both in Table 6.7 and 6.8, one can verify that the ones provided by the heuristic method are very similar to the optimization software ones. In fact, the differences between the results obtained by both resolution methodologies are significantly low. More specifically, for the small instances the differences were always below 2% and for the medium size instances the differences are below 20%.

In terms of computational time, for the instances that GAMS software was able to find a solution, the heuristic multi-start algorithm found the solution in much less time than GAMS. The heuristic method was also able to provide results when the optimization algorithm could not, as one can see from the larger instances and instance 7.

By looking in detail to the results obtained for the values of $\theta = 50$ and $\omega = 60$, the parameter values that were found to suit best the entity's purpose, it was also possible to note that they outperformed the other parameters.

In fact, for the penalization factor of $\theta = 50$ the objective function differences were all lower than the ones obtained by the penalization factor of $\theta = 25$. With regard to when $\theta = 75$, while for the small size instances this parameter had lower differences when comparing to $\theta = 50$, for the medium size the opposite occurred. In terms of the total time and the number of boxes arriving in the laboratory in the busiest period, when the heuristic method provided different solutions between the penalization values, as it is the case of instances 9, 11 and 14, the routes' total time was always less for the penalization factor of $\theta = 50$, supporting the decision of choosing this value. In terms of computational time, the differences between the penalizations were not significant.

By looking into the results of Table 6.8, for the discretization of time into periods of 60 minutes, $\omega = 60$, in terms of the computational time this parameter outperformed significantly $\omega = 30$ since the time taken to obtain the solution was remarkably lower, except for instance 13 and 14 where the opposite occurred. In terms of the objective function, the results were the same in the small and medium size instances. For the larger instances, as expected, the values of the objective function were lower for $\omega = 30$ since that if the day is divided into periods of 30 minutes the number of intervals for the arrivals is the double of when the day is discretized into periods of 60 minutes. By looking into the routes' total time, one can encounter different performances, since that for some instances the results for $\omega = 30$ outperformed the ones for $\omega = 60$, while for others the opposite occurred.

The last analysis lied on finding what is the most suitable value of iterations that balances the computational time and the quality of the solution. Henceforth, tests were performed for a different number of iterations, namely 1 iteration, 50 iterations and 100 iterations. These results are compared to the ones obtained with the optimization program. Table 6.9 depicts the results by presenting the value of the objective function, the number of boxes arriving in the busiest period (w), the difference in percentage (i.e., the relative error when comparing to the results obtained by GAMS) and the computational time (in seconds).

Table 6.9 – Heuristic results for different number of iterations considering $\omega = 60$ and $\theta = 50$.

#	Best GAMS		Heuristic (1 iter.)				Heuristic (50 iter.)				Heuristic (100 iter.)			
	Objective Function	w	Objective Function	w	Difference (%)	Computational Time (sec)	Objective Function	w	Difference (%)	Computational Time (sec)	Objective Function	w	Difference (%)	Computational Time (sec)
1	366	5	368	5	0.5	0,5	368	5	0.5	21	368	5	0.5	39
2	663	10	670	10	1.1	5	670	10	1.1	241	670	10	1.1	449
3	841	12	841	12	0	4	841	12	0	438	841	12	0	840
4	941	15	942	15	0.1	26	942	15	0.1	1331	942	15	0.1	2440
5	952	15	955	15	0.3	16	955	15	0.3	767	955	15	0.3	1392
6	1264	17	1267	17	0.2	55	1267	17	0.2	2432	1267	17	0.2	4410
7	*	*	1199	16	-	29	1199	16	-	1088	1199	16	-	1991
8	790	10	1008	13	27.6	11	800	10	1.3	550	800	10	1.3	1026
9	1061	14	1098	15	3.5	17	1098	15	3.5	1000	1098	15	3.5	1834
10	1114	15	1287	17	15.5	25	1287	17	15.5	1307	1287	17	15.5	2341
11	*	*	1740	19	-	187	1603	17	-	8920	1603	17	-	15936
12	*	*	1683	19	-	228	1623	17	-	10352	1623	17	-	19131
13	*	*	2699	31	-	557	2699	31	-	28279	2699	31	-	56126
14	*	*	2646	12	-	235	2646	12	-	10650	2640	13	-	23148
15	*	*	1685	20	-	134	1685	20	-	7058	1685	20	-	14147

As can be observed, as the number of iterations increases the time taken to obtain the solution also increases. As expected, the number of iterations is in direct proportion to the computational time, since that the time taken to obtain the results for 100 iterations takes approximately 100 times more than for 1 iteration.

In terms of the quality of the results obtained, considering that the objective is to minimize the value of the objective function, it is possible to perceive that the only number of iterations that provided remarkably worse results was for 1 iteration. While for most instances all three values of iterations granted the same results, for instances 8, 11 and 12, the value of 1 iteration provided worse results. Furthermore, with the results obtained for 50 and 100 iterations, one can perceive that the results were the same except for instance 14 where with 100 iterations the program was able to find a better solution.

In conclusion, the quality of the results having several iterations of 100 is better than for the other two parameters. Nevertheless, if one is working with large sized instances, 100 iterations can take too much computational time to be completed. Thus, if the aim is to work with a relatively small amount of collection requests 100 iterations should be the value chosen, while if one is working with very large instances a smaller number of iterations must be selected. Therefore, for the purpose of this thesis, in the next chapter the number of iterations considered is 100 iterations for when dealing with instances with a smaller number of collection requests and 50 iterations for when dealing with many collection requests.

6.4 Chapter conclusions

In summary, the heuristic algorithm is a good alternative to the optimization solver provided with GAMS. While GAMS was only able to find results for small and medium size instances the multi-start algorithm was able to find results for all the instances. In terms of the parameters to use, from the performed tests it can be concluded that the values that fit best the purpose of the routes of Region A are to have a penalization factor of 50 ($\theta = 50$) and a time discretization into periods of 60 minutes ($\omega = 60$). On the other hand, for the purpose of the routes in Region B, since the objective is not to desynchronize the arrivals a penalization of 0 should be considered. Moreover, and similar to Region A, the parameter that discretizes time for Region B should be $\omega = 60$. Finally, the number of iterations to consider should be 50 for when the number of collection requests is very high and 100 when one deals with a smaller number of SCCs. Thus, these are the values considered when using the heuristic methodology to find routes in the following chapter.

7 Case Study Results and Discussion

The aim of this chapter lies on constructing routes for the context of the case study previously outlined in Chapter 2. For this, the multi-start heuristic algorithm for the BSTP with desynchronized arrivals and open routes, already developed and tested, is used to calculate routes for Region A and for Region B.

For this purpose, the parameters that provided the best results in the previous chapter are utilized. More specifically, for both regions a time discretization into periods of 60 minutes, $\omega = 60$, is used. For Region A, since the aim is to desynchronize the arrivals a penalization factor of $\theta = 50$ is employed. On the other hand, for Region B, since the aim is to minimize the routes total time a penalization of $\theta = 0$ is considered.

This chapter is divided as follows: section 7.1 details for Region A a description of the current routes, of the routes provided by the heuristic algorithm and a comparison between them. Section 7.2, on the other hand, details for Region B the currently established routes, the proposed routes and a comparison with each other. On section 7.3 recommendations for the entity are detailed. Finally, section 7.4 presents a final conclusion on the results obtained.

7.1 Region A routes

As already mentioned, associated with the manner that Region A routes are currently defined arises the issue of having periods of the day where too many samples arrive in the laboratory and others where very few arrive. This leads to significant differences in the laboratory throughput throughout the day. Thus, emerging the need of improving these routes by desynchronizing the arrivals in the laboratory.

The following sections have detailed information regarding the routes. Particularly, section 7.1.1 presents an overview of the currently established routes, section 7.1.2 describes the results obtained with the heuristic algorithm and, finally, in section 7.1.3, a comparison of the results with the currently established routes is performed.

7.1.1 Region A current routes

Since the total number of collection requests in Region A is relatively high (more specifically, there are currently 449 requests) designing routes considering all of these locations at the same time is very complex and time consuming to do by hand. Therefore, to overcome this challenge, the entity follows the strategy of dividing Region A into smaller regions based on geographical locations and designs the routes individually for each smaller region. As mentioned in the previous chapter, the names of these regions are: Region A1, Region A2, Region A3, Region A4, Region A5, Region A6 and Region A7.

The entity currently counts on four different types of routes defined by their corresponding schedules. These are:

- Routes that occur in the early morning, more specifically until 12h, which allow samples to arrive early to the laboratory;
- Routes defined in the morning period that arrive in the laboratory at lunch time;
- Routes that collect samples in the afternoon period;
- Routes that visit more distant points that are defined for the full day.

Thus, to a driver, it can be either be attributed an early morning route along with a morning route and an afternoon route, or a morning route and an afternoon route, or a route that takes the full day to be completed. In total there are 39 routes currently established.

Further, having the order of the SCCs to be visited defined, with the purpose of analysing the numerical properties of the routes and the arrivals to the laboratory, it is necessary to have information on the departing times as well as on the time that it takes to travel between locations. Thereby, the entity under study provided information regarding the time at which each route starts. To know the time the routes would take, the time matrix that had to be plotted to calculate the new routes was used. Moreover, the loading time at each SCCs had to be considered, and as experts stated, it takes the value of 10 minutes. Further, to examine the routes in terms of the total travelled distance, the distance matrix that had to be calculated for the new routes was also used for this purpose.

The results obtained were that in total 3744 kms are travelled on each day, taking a total time of 7642 minutes. This is, in average, 117 kilometres and 239 minutes per route. The shortest route has a total distance of 25 kilometres and takes 85 minutes to be completed. The longest, on the other hand, has a total distance of 522 kilometres and takes 517 minutes to be completed. This information is summarized in the next table, Table 7.1.

Table 7.1 – Information regarding the current routes of the entity in Region A.

	Distance (kms)	**Travel time (min)**
Total	3744	7642
Average	117	239
Minimum	25	85
Maximum	522	517

After having the schedules of the routes, it was possible to compute a graph with the predictions of what the arrivals in the laboratory currently are. Hence, the graph in Figure 7.1 was plotted. This graph represents the number of boxes arriving in the laboratory (i.e., the number of SCCs visited by the given routes) in function of the corresponding arrival times (divided into intervals of 1 hour).

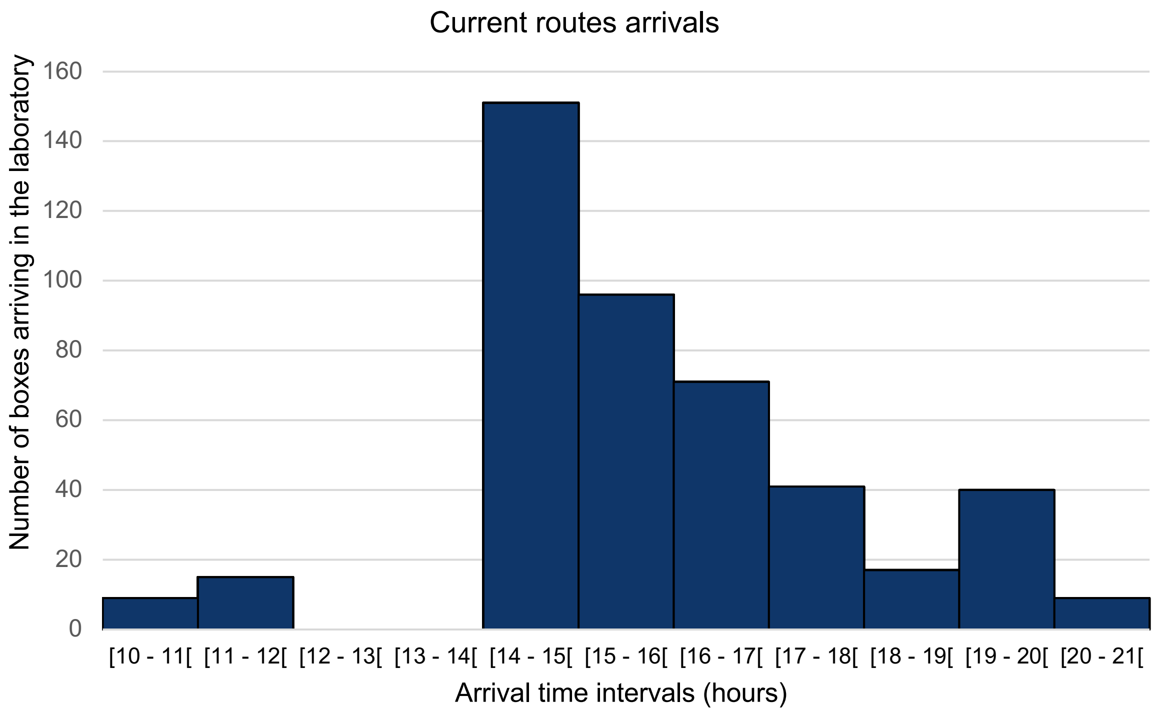


Figure 7.1 – Estimation of the number of boxes arriving in the entity's central laboratory per hour by the routes organized in Region A.

Analysing Figure 7.1, one can comprehend the flow of the arrivals in the laboratory. Firstly, the impact of the early morning routes can be perceived by the samples that arrive between 10h and 12h. In the interval of time from 12h to 14h there are not any samples arriving in the laboratory. In the intervals after is where most of the samples arrive, more precisely, in the time period from 14h-16h, reaching the peak of 151 boxes between 14h-15h. These arrivals mainly come from the samples collected by the morning routes. The arrivals of the samples after 16h are more distributed in time and they mainly come from the afternoon routes and longer routes. This disparity in the arrivals, where there is a period where no samples arrive and another where most of them arrive, represents the main issue faced in Region A routes, which portraits the need of rearranging the routes.

Finally, having the current routes described it is possible to understand the need of desynchronizing the routes arrivals in order to allow for the laboratory to have a more continuous flow of samples arrivals. Thus, the objective is to minimize the peak that occurs in the interval from 14-15h, while at the same time try to minimize the total time of the routes.

7.1.2 Region A proposed routes

To calculate the new routes for Region A, experts stated that the early morning routes should not be modified. Hence, only the morning routes, afternoon routes and the routes that take the full day to be completed are considered. For these, there is a total number of 425 collection requests.

As previously detailed in Chapter 6, when calculating routes using the heuristic algorithm, as the number of requests increase the computational time taken to compute the routes also increases. Hence, considering all the 425 request nodes at the same time would not be adequate since it would take a considerable amount of time to obtain results. In order to overcome this challenge, a strategy of diving the request nodes into different clusters based on geographical locations was followed. Firstly, the request nodes from Region A6 and A7 were considered in a different group from Region A1, A2, A3, A4 and A5.

The routes presented in Region A6 and A7 are the ones that take the full day to be completed, having a total of 60 collection requests. According to entity experts, different input parameters must be considered for these locations, namely the average velocity to calculate the time matrix and the maximum time limitation for the routes. For this, an average velocity of 75 km/h and a maximal time for each route of 540 minutes should be considered.

Even with the two geographical clusters presented, one would still have 365 request nodes for the regions of Region A1, A2, A3, A4 and A5. Therefore, since this number is notably high, it would still not be possible to obtain results within reasonable computing time. Since the aim is to desynchronize the arrivals, dividing the cluster into smaller ones based on geographical locations is not an option due to the way the mathematical model was developed. In fact, the arrivals would not be desynchronized, since it would only consider the desynchronization of each individual region. Therefore, the strategy followed was to divide the request nodes into two different groups based on their collecting schedules. Therefore, one group is composed by the request nodes of the morning routes (250 request nodes) and the other consists of the nodes in the afternoon routes (115 nodes). According to the entity experts, an average velocity of 60 km/h and a total travel time limitation of 360 minutes should be considered.

Hence, the strategy taken was to first divide the request nodes into two different clusters based on geographical locations, and then divide the request nodes of the largest cluster into two groups based on the collecting schedules. A representation of the strategy followed is represented in the following image, Figure 7.2.

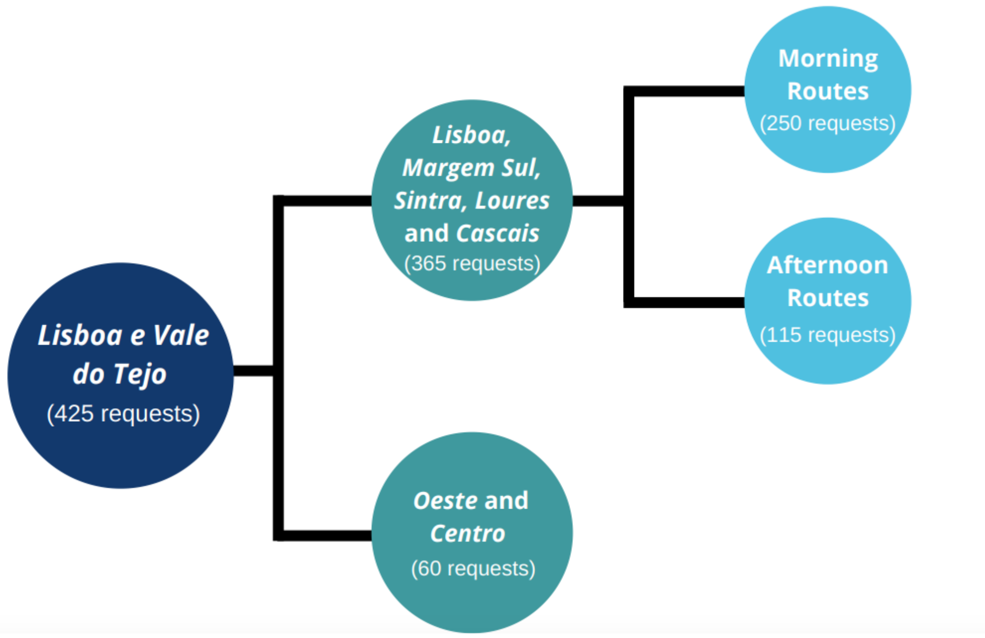


Figure 7.2 – Representation of the strategy taken to divide the total number of requests to calculate new routes.

Apart from the maximal time each route should take and the average velocity to calculate the time matrix, other input parameters must be considered. Similar to the instances created in the previous chapter, there are common input parameters to be introduced. Furthermore, as stated, for all the groups a penalization of $\theta = 50$ and a time discretization into periods of 60 minutes, $\omega = 60$, was considered. Regarding the number of iterations, for Regions A6 and A7 100 iterations were considered. For the morning routes of Region A1, A2, A3, A4 and A5, since this group deals with the greatest number of request nodes 50 iterations were chosen. Finally, for the afternoon routes 100 iterations were considered. All the presented input parameters are summarized in Table 7.2.

Table 7.2 – Input parameters considered for each zone in Region A.

Group	# Collection Requests	Number of iterations	θ	ω (min)	T_{max} [1] (min)	τ_0 [2] (min)	τ_i [3] (min)	a_0 [4] (min)	b_0 [5] (min)	M [6]
Region A6 and A7	60	100	50	60	540	-	10	480	1260	100000
Morning routes	250	50	50	60	360	-	10	480	1260	100000
Afternoon routes	115	100	50	60	360	-	10	480	1260	100000

[1] T_{max} is the maximal transportation time for each route
[2] τ_0 is the unloading time in the laboratory. Since only one route is considered for each vehicle, this value is not considered when calculating these routes
[3] τ_i is the unloading time in each SCC
[4] a_0 represents the upper time window of the laboratory
[5] b_0 represents the lower time window of the laboratory
[6] M is a large constant

After introducing the input parameters into the algorithm, for each region new routes were retrieved. The numerical results obtained were the following:

- For the routes of Regions A6 and A7:
  - The routes' total time would be 2074 minutes and the total distance 1843 kms;
  - Comparing to the current routes it would take more 61 minutes and 78 kms to be performed.
- For the morning routes of Region A1, A2, A3, A4 and A5*:*
  - The routes' total time would have a value of 3754 minutes and the total number of kilometres would be 1254;
  - Comparing to the current routes it would be necessary to travel more 97 minutes and 97 kms.
- Finally, for the results obtained for the afternoon routes of Region A1, A2, A3, A4, and A5*:*
  - The routes' total travel time would be of 2011 minutes and the total number of kilometres 861;
  - Comparing to the current solution it would take more 39 minutes and 39 kms.

In summary, the routes obtained with the algorithm sum in total 3958 kms travelled within 7839 minutes, which is more 214 kms and 197 minutes than the current solution. This is in average 124 kms and 245 minutes per route. The shortest route would have a total distance of 37 kms and take 97 minutes to be completed. The longest, on the other hand, would have a total distance of 536 kms and take 485 minutes to be completed. Table 7.3 summarizes all the presented numerical results for the routes obtained with the multi-start heuristic algorithm.

Table 7.3 – Information regarding the routes obtained with the multi-start algorithm in Region A.

	Distance (kms)	**Travel time (min)**
Total	3958	7839
Average	124	245
Shortest route	37	97
Longest route	536	485

Additionally, and to comprehend the distribution of the arrivals during the day, a graph like the one presented in the previous section (Figure 7.1) was developed. Thus, the next figure, Figure 7.3, represents this graph which depicts the number of boxes arriving in the laboratory (i.e., the number of SCCs visited by the given routes) in function of the arrival time intervals (divided into intervals of 1 hour).

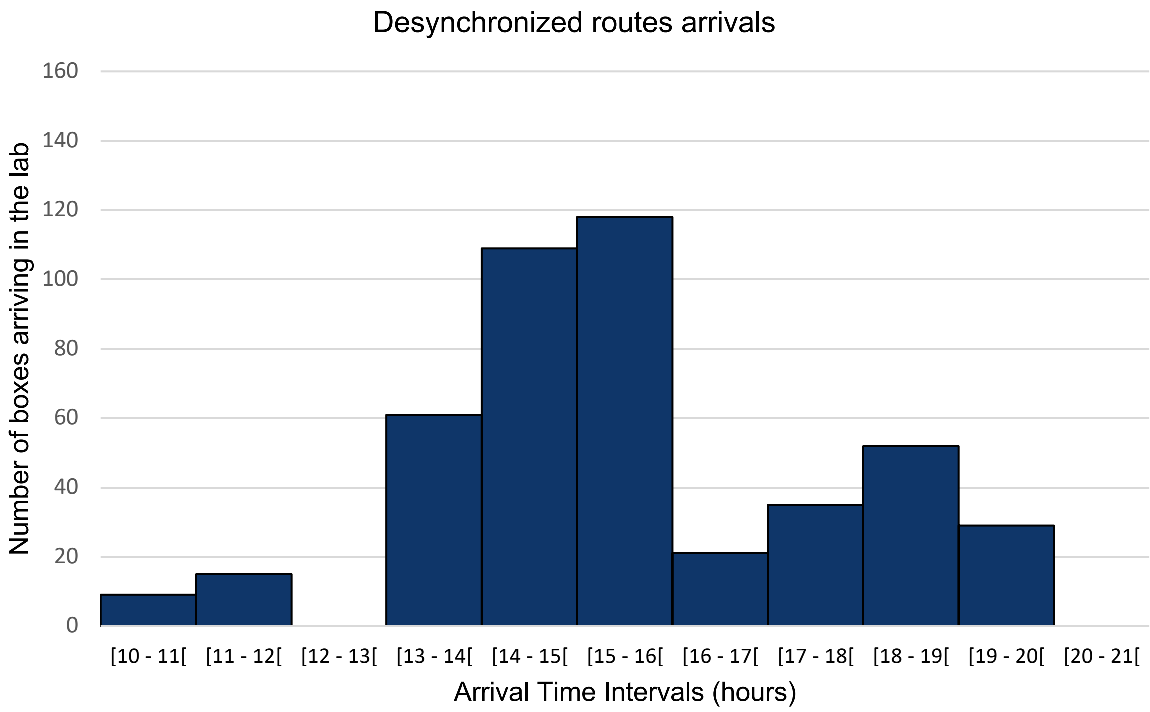


Figure 7.3 – Estimation of the number of boxes arriving in the central laboratory per hour by the routes organized in Region A.

Analysing Figure 7.3, one can understand what the flow of the arrivals in the laboratory would be using the routes obtained with the heuristic algorithm. Firstly, as expected the arrivals between 10h and 12h would be the same, since the early morning routes were not modified. After this period, one would not have arrivals between 12h to 13h, having samples arriving in the laboratory after 13h. The peak of the maximum number of arrivals would be reached between the time period of 15h-16h reaching 118 boxes, less 33 boxes than the busiest period of the current routes. After 16h one would have less arrivals coming mainly from the afternoon routes and from the routes that take the full day to be established.

7.1.3 Region A routes' comparison

Having described in detail both the current routes and the routes obtained with the heuristic algorithm, it is possible to perceive the main differences between them. In fact, the results obtained with the heuristic method were improved in terms of the desynchronization of the arrivals in the central laboratory; nevertheless, desynchronizing the arrivals implies an increase in the routes total travelled time and consequently in the total distance travelled. All these differences are now detailed.

Firstly, a comparison on the numerical results obtained for both the current routes and for the proposed routes was performed. The results are represented in the following table, Table 7.4.

Table 7.4 – Numerical comparison of the routes currently established with the ones provided by the heuristic method in Region A.

	Current routes		Proposed routes	
	Distance (kms)	Travel time (min)	Distance (kms)	Travel time (min)
Total	3744	7642	3958	7839
Average	117	239	124	245
Minimum	25	85	37	97
Maximum	522	517	536	485
Number of boxes arriving in the busiest time period (w)	151		118	

Analysing Table 7.4, it is possible to understand that there is an increase of 214 km and of 197 minutes on each day. In percentage, this corresponds to an increase of 5.7% in terms of the total distance and of 2.6% in terms of time. The routes in average would also take more 6 minutes and 7 kilometres to be performed. Comparing the results for the shortest and longest routes, one would have more homogeneous routes with the ones provided by the heuristic algorithm since that the shortest route would take more time to be completed than the one that is currently established, and for the largest, one the other hand, it would take less time to be performed. Moreover, the fact that the longest route could be performed in less time is also beneficial for the entity due to the perishable nature of biomedical samples. Regarding the highest number of boxes arriving in the busiest time period, a decrease of 33 boxes would be attained.

Secondly, concerning results obtained for the arrivals of the current ones (presented in Figure 7.1) and the proposed ones (presented in Figure 7.3) major differences can be identified. More specifically, apart from the early morning routes where the arrivals are maintained, with the new routes it would be possible to have boxes arriving earlier in the laboratory. As a matter of fact, one would have samples arriving at 13h in comparison to the current routes where samples only arrive after 14h. Also, the peak of 151 boxes that occurred from 14h to 15h in the current routes would pass to occur between 15h and 16h having 118 boxes arriving in that interval, a difference of 33 boxes. Another important fact to point out lies on the afternoon peak that occurs in the current routes in the interval from 19h to 20h that would occur earlier for the proposed routes, between 18h and 19h. Finally, the routes obtained with the metaheuristic would also allow not having boxes arriving in the laboratory after 20h which is beneficial since it could permit the laboratory to close earlier. In general, the new routes would allow having samples arriving earlier, providing for a smoother workflow in the laboratory. Further, allowing the samples to be analysed earlier which in turn would make results to be sent to patients also earlier.

Finally, to further analyse the impact of these arrivals in the workflow of the laboratory, and since the labs' capacity is not explicitly known in terms of how many boxes can be analysed per hour, three different scenarios were created to analyse different possible capacities for the laboratory workflow. The three scenarios lie on having a laboratory capacity of 60, 80 or 100 boxes per hour.

Scenario 1 – Labs' capacity of 60 boxes per hour

Figure 7.4 depicts the laboratory capacity usage in percentage in terms of the working time in the laboratory if the total capacity is 60 boxes per hour. Namely, Figure 7.4a represents it for the current established routes and Figure 7.4b for the routes obtained with the heuristic algorithm.

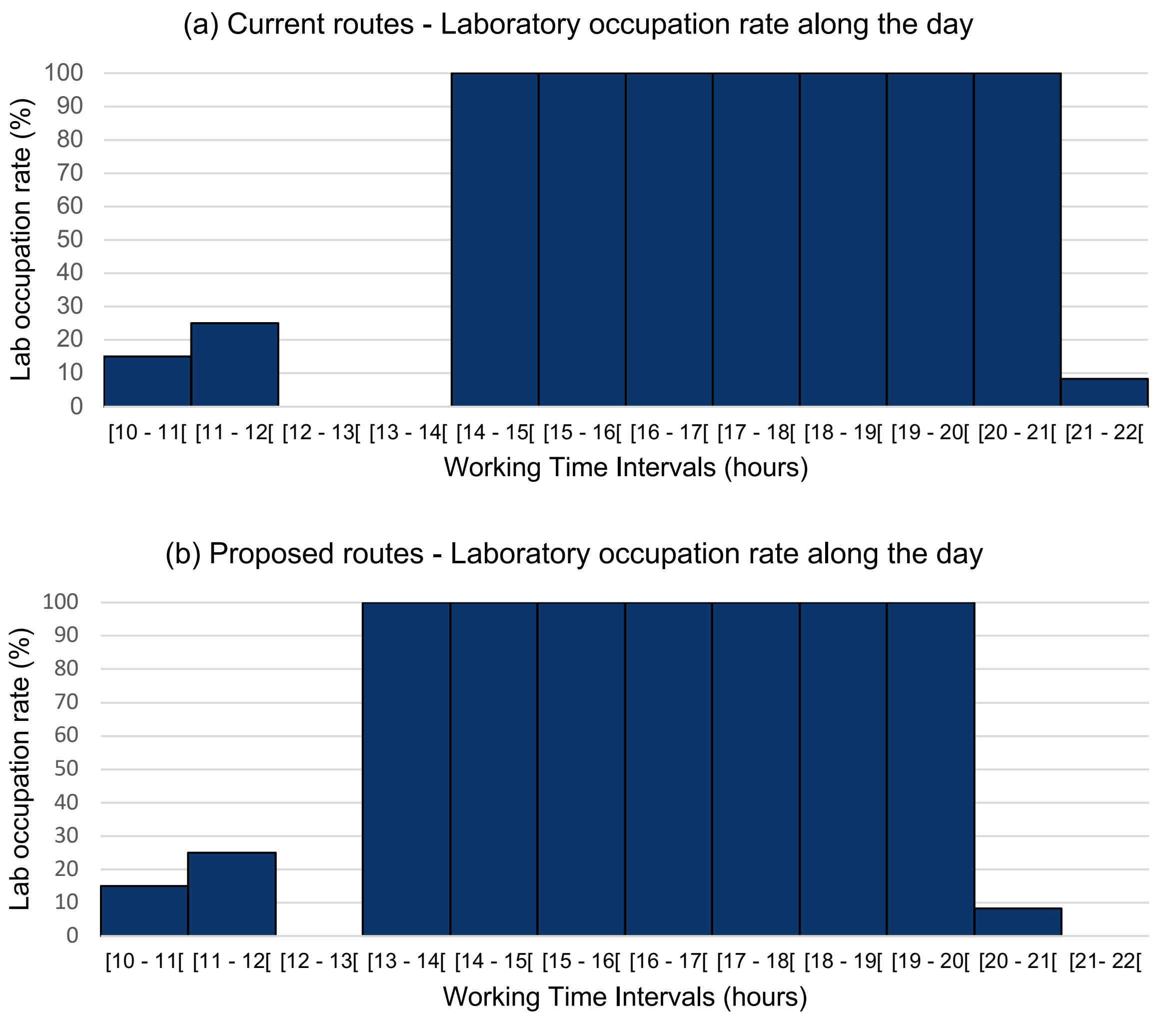
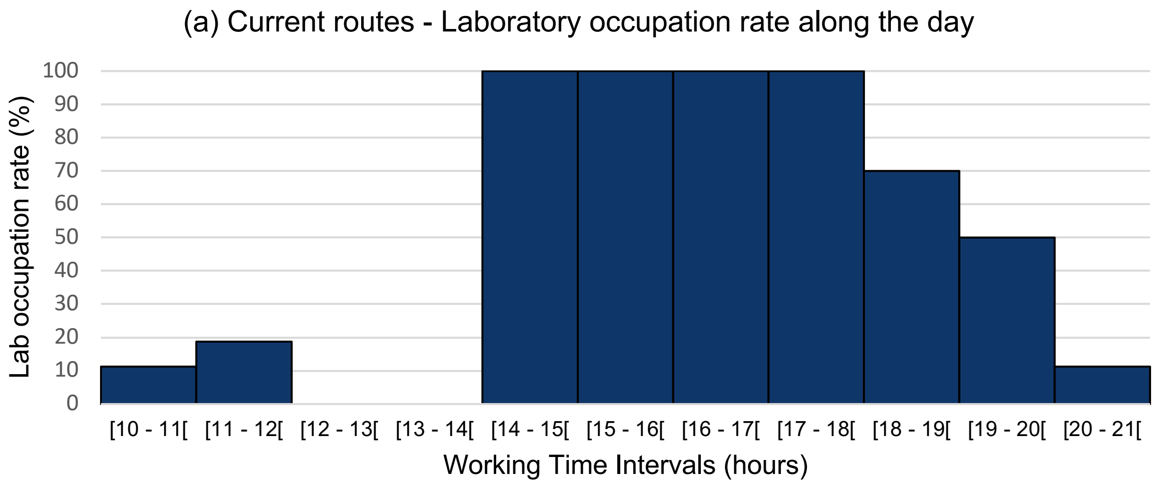


Figure 7.4 – Laboratory occupation rate in percentage through the day considering a total capacity of 60 sample boxes per hour. **Figure 7.4a:** Current Routes; **Figure 7.4b:** Proposed routes.

Comparing both graphs, it is possible to understand that the capacity of the laboratory in the new routes would start to be fully utilized since 13h on the contrary o the current routes where it only starts to be utilized after 14h. In this scenario, for the routes obtained with the heuristic algorithm it would also enable all the samples to be analysed one hour earlier.

Scenario 2 – Labs' capacity of 80 boxes per hour

For the second scenario, a labs' total capacity of 80 boxes per hour is tested. The results obtained for the laboratory occupation rate throughout the day for both the current routes and for the proposed routes are represented, correspondingly, in the following figures, Figure 7.5a and 7.5b.



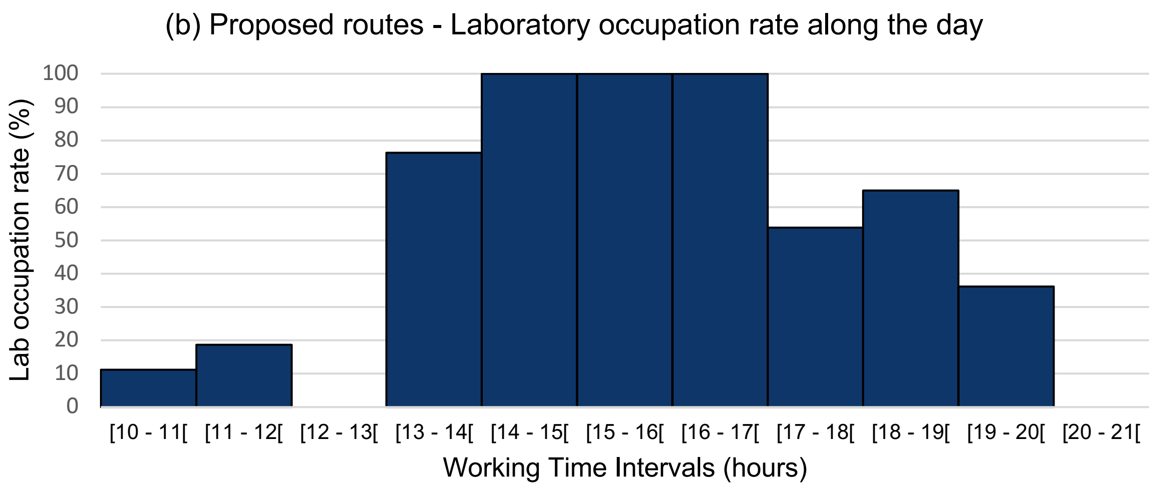
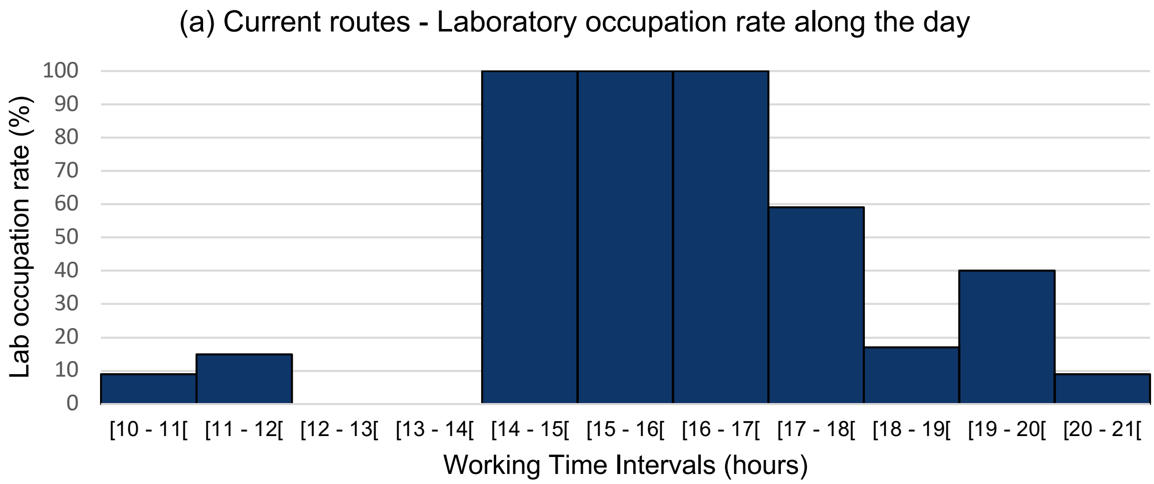


Figure 7.5 – Laboratory occupation rate in percentage through the day considering a total capacity of 80 sample boxes per hour. **Figure 7.5a:** Current Routes; **Figure 7.5b:** Proposed routes.

Comparing both graphs, it is possible to perceive that with the proposed routes all the samples would, once again, start being analysed one hour earlier. Consequently, the possibility of having all of them analysed one hour earlier arises. Further, for the proposed routes one would only have the full capacity being used for 3 hours, on the contrary to the current routes where it is utilized for 4 hours, allowing for a smoother workload.

Scenario 3 – Labs' capacity of 100 boxes per hour

Finally, for the last scenario, 100 boxes per hour, the results obtained for the laboratory occupation rate through the day are presented in Figure 7.6. More specifically, the results for the current routes are described in in Figure 7.6a and in Figure 7.6b for the proposed routes.



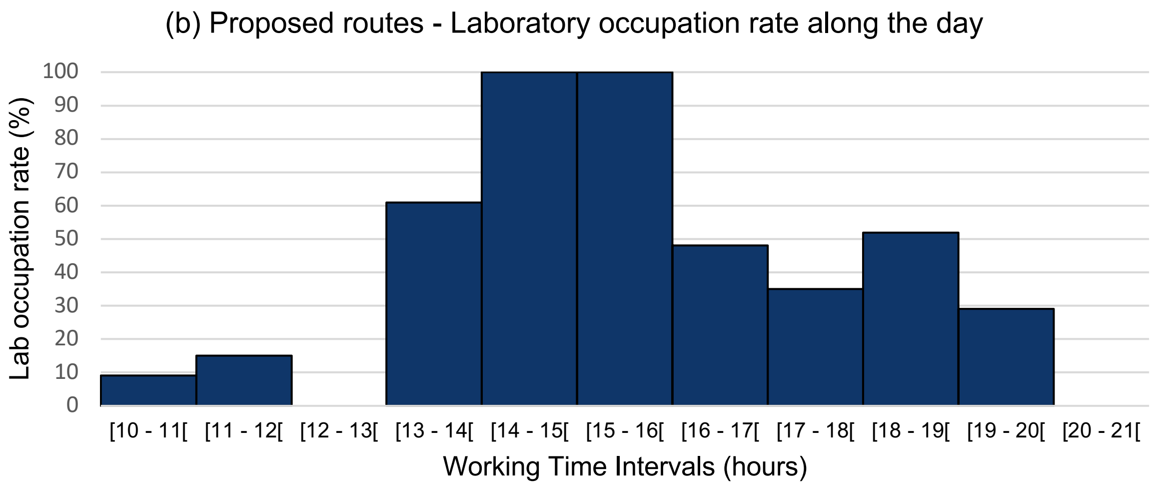


Figure 7.6 – Laboratory occupation rate in percentage through the day considering a total capacity of 100 sample boxes per hour. **Figure 7.6a:** Current Routes; **Figure 7.6b:** Proposed routes.

Finally, for the last scenario, if one would have a total capacity of analysing 100 sample boxes per hour, once again, all the samples would be analysed 1 hour earlier with the proposed routes. Furthermore, for the proposed routes the laboratory occupation rate would be fully used for 2 hours while for the current routes it would be for 3 hours.

With all being said, for all the three different scenarios it would be possible to have all the samples analysed one hour earlier comparing to the currently established routes. Also, for scenario 2 and 3 it would allow having the full capacity being utilized for less hours, providing for a smoother workload in the laboratory.

In summary, although the proposed routes require an increase in terms of the routes total distance and total travelled time, it would be possible to have all the laboratory work to be completed earlier. Not only this would allow results to be sent earlier to patients, but also would allow freeing the technicians that had to work during this hour, which could in turn compensate the cost associated to the increase in routes total travelled time and distance.

7.2 Region B routes

For the routes established in Region B that transfer biomedical samples to the central laboratory, the issue of having too many samples arriving in the laboratory in certain periods of the day is not a concern. The main challenge that is faced in this region lies on the fact that the routes take several hours to be completed, which makes these samples arrive very late to the laboratory to be analysed.

The following sections have detailed information regarding the routes in Region B. Specifically, section 7.2.1 outlines the currently established routes, section 7.2.2 describes the results obtained with the heuristic algorithm and, finally, in section 7.2.3, a comparison of the current routes with the ones provided by the algorithm is established.

7.2.1 Region B current routes

In Region B, the entity has currently three pre-defined routes established, which occur in every weekday to collect samples from 26 different SCCs. All these routes depart in the morning from one of the entity's laboratories that is located in Region B, visit the corresponding SCCs and terminate once they deliver the biomedical samples in the central laboratory.

Then, to analyse in detail these routes, the time and the distance that they take to be completed was estimated. To estimate it and to obtain values that can be compared to the ones that the algorithm provides, once again, both the distance matrix and the time matrix that had to be computed to calculate the new routes were used. Thus, with the information of the departing time of each route, the time the routes would take was estimated. It is important to mention that the 10-minute interval stated that takes to load the samples in the SCCs was considered. Finally, the drivers' 1-hour lunch break was also taken into consideration.

The results obtained were the following:

- The first route is scheduled to occur from 10h00 to 19h01, taking around 9 hours and 502 km to be completed;
- The second one is scheduled to occur from 11h00 to 18h04, taking around 7 hours and 355 km to be completed;
- Finally, the last one is scheduled to occur from 10h00 to 20h42, taking almost 11 hours and 590 km to be completed.

This gives the routes of Region B a total duration of 26h47 which is in average 09h01 per route. In terms of distance, the drivers travel a total distance of 1447 km which is in average 482 kms per route.

Finally, as one can verify from these results, the samples are arriving very late in the laboratory, especially for the last route that takes almost 11 hours to be completed and is arriving at almost 21h00. Thus, arising the need of rescheduling these routes.

7.2.2 Region B proposed routes

The aim when proposing the new routes for Region B is to minimize the routes' total travel time, hence the multi-start heuristic algorithm is used with the parameters that were found in Chapter 4 to fit best the case study purpose for this region. In fact, a penalization of 0 (i.e., $\theta = 0$) and a time discretization into periods of 60 minutes (i.e., $\omega = 60$) are considered. Moreover, since the number of collection requests is not very high (there are 26 collection requests) a total number of 100 iterations is used.

Furthermore, when discussing with entity experts what should the time limitation for the drivers be, it was stated that a time limitation of 09h30 per driver (including the lunch time) is appropriate. With this value, one would have routes that are fairer for the drivers in terms of their working time and not have routes that take too much time to be completed as the ones that are currently established.

Regarding the average time to compute the time matrix, experts stated that an average velocity of 75 km/h should be considered. Furthermore, the average waiting time in the SCCs should be 10 minutes and the laboratory working hours should be the same as previously considered.

All the presented parameters are now schematized in the following table, Table 7.5.

Table 7.5 – Input parameters considered for Region B.

# Collection Requests	Number of iterations	$\boldsymbol{\theta}$	$\boldsymbol{\omega}$ (min)	$\boldsymbol{T_{max}}$ [7] (min)	$\boldsymbol{\tau_0}$ [8] (min)	$\boldsymbol{\tau_i}$ [9] (min)	$\boldsymbol{a_0}$ [10] (min)	$\boldsymbol{b_0}$ [11] (min)	$\boldsymbol{M}$ [12]
26	100	0	60	570	-	10	480	1260	100000

Subsequently, after inserting the input parameters into the metaheuristic, new routes were retrieved. The numerical results obtained were the following:

- The first route would be scheduled to occur from 11h43 to 18h39, taking around 7 hours and 383 km to be completed;
- The second route would be scheduled to occur from 10h02 to 18h09, taking around 8 hours and 355 km to be completed;
- Finally, the last one would be scheduled to occur from 09h56 to 18h58, taking around 9 hours and 590 km to be completed.

These routes sum in total 24h05 which is in average 08h07 per route. In terms of distance, the drivers would have to travel a total distance of 1245 km which is 415 kms per route.

7.2.3 Region B routes' comparison

From the results obtained, one can corroborate that the routes provided by the multi-start heuristic method were improved in comparison to the current ones, both in terms of the routes' total time and in terms of the travelled distance. A numerical comparison for the current routes and for the ones obtained with the heuristic algorithm is now detailed in Table 7.6.

[7] T_{max} is the maximal transportation time for each route
[8] τ_0 is the unloading time in the laboratory. Since only one route is considered for each vehicle, this value is not considered when calculating these routes
[9] τ_i is the unloading time in each SCC
[10] a_0 represents the upper time window of the laboratory
[11] b_0 represents the lower time window of the laboratory
[12] M is a large constant

Table 7.6 – Numerical comparison of the routes currently established with the proposed routes provided by the heuristic method.

		Current Routes	Proposed Routes	Improvement
Time	Route 1:	10h00–19h01 (dur. 09h01)	11h43–18h39 (dur. 06h56)	-
	Route 2:	11h00–18h04 (dur. 07h04)	10h02–18h09 (dur. 08h07)	-
	Route 3:	10h00–20h42 (dur. 10h42)	09h56–18h58 (dur. 09h02)	-
	Total time:	26h47 (1607 min)	24h05 (1445 min)	**02h42**
	Average:	~ 09h01/route	~ 08h07/route	**54 min/route**
Distance	Route 1:	502 kms	383 kms	-
	Route 2:	355 kms	397 kms	-
	Route 3:	590 kms	465 kms	-
	Total distance:	1447 kms	1245 kms	**202 kms**
	Average	~ 482 kms	~ 415 kms	**67 kms/route**

By analysing Table 7.6, one can observe the improvements in terms of time and distance. In fact, with the proposed routes an improvement of 02h42 (i.e., less 11.2% in terms of time) would be achieved. Regarding the total distance, an improvement of 202 kms (i.e., less 16.2% in terms of distance) would be attained. Concerning the schedules of the routes the maximum time that a route would take would pass from 10h42 to 09h02. The shortest route would also diminish since it would pass from taking 07h04 to be around 06h56. The remaining route would also be improved since it would change to last from 09h01 to last 08h07. As a matter of fact, the results provided by the heuristic algorithm provide more balanced routes in terms of time since the difference between the longest and the shortest route is less than from the ones currently established.

Furthermore, the routes retrieved by the metaheuristic were also improved regarding the arrival times, since all of them would arrive before 19h on the contrary to the currently established ones where the longest route would arrive at almost 21h. This would enable samples to arrive earlier at the laboratory and to be tested earlier which in turn would allow results to be sent to the patients also earlier.

In conclusion, the heuristic algorithm was very effective calculating Region B routes since it was able to provide the results desired by the entity. Not only the routes were improved in terms of distance and total travel time, but it would also enable results to be delivered earlier to patients.

7.3 General recommendations

From the previous discussion, one can conclude:

- The results obtained for the Region A show that there are changes that can be performed to smooth the workload in the laboratory. Even though the program was not able to reduce the total travelled time (having an increase of 5.7% in terms of distance and of 2.6% in terms of time) with the results obtained it is possible to have earlier arrivals in the laboratory and have all the samples analysed one hour earlier, and consequently deliver earlier the results to patients;
- With the current resources, the expected improvement in terms of kilometres for the routes of Region B is of 16.2% and of 11.2% in terms of the total travelled time;
- The results for Region B also exhibit the possibility of having earlier results deliveries to patients. Having earlier arrivals in the laboratory would enable samples to be analysed earlier which in turn would allow earlier results deliveries to patients.

7.4 Chapter conclusions

In this chapter the heuristic algorithm was applied to the case study previously presented in Chapter 2. Thus, new routes were created for Regions A and B. While the aim for the creation of new routes for Region A was to desynchronize the arrivals taking into account the minimization of the routes total time, for Region B the only goal was to minimize the routes total travel time.

For Region A, to desynchronize the arrivals a penalization of $\theta = 50$ was considered. Since the number of nodes in this region is very high, a strategy of dividing them into different groups and compute new routes for each one of these groups was performed. Firstly, they were divided into two groups based on geographical locations having the collection requests for Regions A6 and A7 apart from the other regions. Since in the other regions there would still be many nodes, the strategy of dividing them based on their collection schedules was performed; thus, the morning routes were in a different group than the afternoon routes. The results obtained with the heuristic algorithm were favourable in terms of the samples arrivals since they would be able to arrive earlier in the laboratory and to be analysed earlier. Nevertheless, the desynchronization of the arrivals implies an increment of 5.7% in the total distance and of 2.6% in the total time.

On the other hand, for Region B, since the number of nodes is relatively small the whole set was considered at the same time. Since the aim of the creation of these routes is only to minimize the total travel time, a penalization of $\theta = 0$ was considered. The results obtained were favourable and in accordance with what was expected from the entity, having a daily improvement of 16.2% in terms of distances and of 11.2% in terms of time. The new routes also permit earlier arrivals of biomedical samples in the laboratory, allowing them to be analysed earlier making the results being delivered to patients also earlier.

8 Conclusions and Future Work

8.1 Conclusions

This dissertation focuses on addressing the challenge found when designing the routes of a clinical analysis group in two different regions: Region A and Region B. Namely, the routes in Region A collect biomedical samples from different SCCs and transport them to the central laboratory to be analysed. The current routes create at the laboratory certain periods of the day where a large number of biomedical samples arrive while there are others where very few samples arrive. This is causing a bottleneck in the samples' supply chain of the central laboratory. On the other hand, the routes in Region B that start in one of the entity's laboratories located in Region B and terminate in the central laboratory are taking a considerable amount of time to be completed, causing the samples to arrive very late in the laboratory. Consequently, the primary goal of the present research was to apply optimization techniques to assist the entity under study planning their routes.

To deal with these challenges, a literature review on the matter was performed. Based on this, a Biomedical Sample Transportation Problem (BSTP) with desynchronized arrivals was developed where the possibility of open routes in order to fit the purpose of the entity was considered. To the best of one's knowledge this is the first work on the BSTP to include this possibility. The problem then consists in minimizing both the routes total travel time and desynchronizing the arrivals in the laboratory.

Afterwards, the developed mathematical model was implemented in an optimization modelling software, GAMS. To test the model, a set of 15 instances characterized by their size (small, medium and large) was created having as a basis the geographical regions that the entity considers when constructing the routes. Regarding the large instances created, the program was not able to find any solution even after several hours of computation. Consequently, since the entity deals with a great number of SCCs the need of developing a heuristic algorithm emerged.

Thus, a metaheuristic algorithm was developed, more specifically, a multi-start algorithm. In order to test the metaheuristic, the same instances were used, and the results were compared to the ones obtained previously with the optimization program. The heuristic proved to be efficient in minimizing both the routes total duration as well as the number of arrivals during the busiest period, since it provided similar results and in less time in relation to the optimization program. In fact, it presented a relative error always below 20%. In addition, it also granted results for the large instances which GAMS could not find results.

The parameters that fit best the case study context were also studied, arriving to the conclusion that the finest penalization to desynchronize arrivals would be of 50 (i.e., $\theta = 50$) and that the finest time discretization would be into periods of 60 minutes (i.e., $\omega = 60$). Finally, the most appropriate number of iterations for the algorithm was investigated. Hence, arriving to the conclusion that if the aim is to work with a relatively small amount of collection requests it is appropriate to use 100 iterations, while if one is working with a set of large instances a smaller number of iterations must be selected to obtain results requiring computational time.

After having the algorithm developed and tested and having the finest parameters studied, the heuristic method was applied to redesign the entity's routes in the two different regions: Region A and Region B.

For Region A the results obtained were able to fulfil their requirements since the arrivals in the laboratory would be more fluid: having the total number of sample boxes arriving in the busiest period to pass from 151 to 118. Also, all the samples would be able to arrive 1 hour earlier to the laboratory in comparison to the currently established routes. Nevertheless, with the new routes comes the cost of 5.7% in the routes total travelled distance and of 2.6% in terms of the routes total travelled time. Furthermore, an analysis on the laboratory occupation was performed and it was possible to observe that the proposed routes would allow for a smoother workload in the laboratory also enabling samples to be processed earlier. This fact could compensate the increased costs in terms of the total travelled distance and total time since that the technicians working in the laboratory could be dispensed one hour earlier.

On the other hand, for Region B the results obtained were very optimistic, having an improvement of 16.2% in terms of distance and of 11.2% in terms of time on each day. Also, the routes obtained would be more balanced in terms of time and would permit the collected samples to arrive earlier to the laboratory to be analysed, which in turn would also allow results to arrive earlier to patients.

In summary, the proposed model and developed heuristic algorithm were able to find results that fulfilled the entity requirements. More precisely, for Region A desynchronized arrivals in the laboratory would be achieved allowing for a smoother workload in the laboratory. For Region B reduction of the routes total travel time would be attained which in turn would allow the biomedical samples to arrive earlier in the laboratory to be analysed.

8.2 Limitations and future work

Within the developed work, although important results were obtained, also some limitations can be identified that can be taken as basis for future work in the problem. Thus, regarding the heuristic algorithm proposed and in order to obtain solutions with higher quality, changes in the algorithm could be performed. Firstly, further randomization should be introduced in the creation of the initial solution at each iteration in the algorithm. Also, in the improvement steps, diversification of the search should be achieved in order to further explore the solution space. An additional suggestion lies on allowing for soft time windows in order to obtain improved solutions.

Still on the developed metaheuristic improvements in the performance of the algorithm should be carried out in order to reduce the computational time taken to calculate the routes.

Furthermore, it would be interesting to develop a Google Maps API to retrieve more accurate time matrices. Although introducing the time matrix obtained from the distance matrix considering an average velocity in the algorithm provides faithful results, it would be more reliable to have into consideration possible traffic information and velocities that consider different types of roads (i.e., highways, rural areas, etc..).

A final suggestion lies on building a user-friendly interface application so that anyone with access can easily calculate new routes. This would be helpful for the entity under study, since their routes are changed frequently. The application should comprise all the input parameters and would also be interesting if it could retrieve a map with the new routes and the stops identified.

References

[1] “Increasing Workforce Productivity in the Diagnostic Laboratory How Can the Lab Keep Up with Demand?,” *Siemens Healthcare Diagnostics Inc*, 2018. .

[2] G. Erdoğan, “An open source Spreadsheet Solver for Vehicle Routing Problems,” *Comput. Oper. Res.*, vol. 84, pp. 62–72, Aug. 2017.

[3] G. B. Dantzig and J. H. Ramser, “The Truck Dispatching Problem,” *Source Manag. Sci.*, vol. 6, no. 1, pp. 80–91, 1959.

[4] M. M. Flood, “The Traveling-Salesman Problem,” *Oper. Res.*, vol. 4, no. 1, pp. 61–75, Feb. 1956.

[5] C. Alabas-Uslu and B. Dengiz, “A self-adaptive local search algorithm for the classical vehicle routing problem,” *Expert Syst. Appl.*, vol. 38, no. 7, pp. 8990–8998, 2011.

[6] F. Daneshzand, “The Vehicle-Routing Problem,” in *Logistics Operations and Management*, Elsevier Inc., 2011, pp. 127–153.

[7] P. Toth and D. Vigo, *The vehicle routing problem*. Philadelphia: Society for Industrial and Applied Mathematics, 2002.

[8] T. Vidal, G. Laporte, and P. Matl, “A concise guide to existing and emerging vehicle routing problem variants,” *Eur. J. Oper. Res.*, vol. 286, no. 2, pp. 401–416, 2019.

[9] A. Mor and M. G. Speranza, “Vehicle routing problems over time: a survey,” *4OR*, vol. 18, no. 2, pp. 129–149, 2020.

[10] F. Arnold and K. Sörensen, “What makes a VRP solution good? The generation of problem-specific knowledge for heuristics,” *Comput. Oper. Res.*, vol. 106, pp. 280–288, Jun. 2019.

[11] B. L. Golden, S. (Subramanian) Raghavan, and E. A. Wasil, *The vehicle routing problem : latest advances and new challenges*. Springer, 2008.

[12] Tadahiko Murata and Ryota Itai, *Enhancing Solution Similarity in Multi-Objective Vehicle Routing Problems with Different Demand Periods.* INTECH Open Access Publisher, 2008.

[13] A. N. Letchford and J. J. Salazar-González, “Projection results for vehicle routing,” *Math. Program.*, vol. 105, no. 2–3, pp. 251–274, 2006.

[14] T. R. P. Ramos, M. I. Gomes, and A. P. B. Póvoa, “Multi-depot vehicle routing problem: a comparative study of alternative formulations,” *Int. J. Logist. Res. Appl.*, vol. 23, no. 2, pp. 103–120, 2020.

[15] M. Mirabi, N. Shokri, and A. Sadeghieh, “Modeling and Solving the Multi-depot Vehicle Routing Problem with Time Window by Considering the Flexible End Depot in Each Route,” *Int. J. Supply Oper. Manag. IJSOM Shokri Sade. Int J Supply Oper Manag.*, vol. 33, no. 33, pp. 1373–1390, 2016.

[16] J. Widuch, “Current and emerging formulations and models of real-life rich vehicle routing problems,” in *Smart Delivery Systems*, Elsevier Inc., 2020, pp. 1–35.

[17] P. Kilby and P. Shaw, “Chapter 23 Vehicle routing,” in *Foundations of Artificial Intelligence*, vol. 2, no. C, Elsevier B.V., 2006, pp. 801–836.

[18] S. N. Kumar and R. Panneerselvam, “A Survey on the Vehicle Routing Problem and Its Variants,” *Intell. Inf. Manag.*, vol. 04, no. 03, pp. 66–74, 2012.

[19] H. Asefi, S. Shahparvari, P. Chettri, and S. Lim, “Variable fleet size and mix VRP with fleet heterogeneity in Integrated Solid Waste Management,” *J. Clean. Prod.*, vol. 230, pp. 1376–1395, 2019.

[20] R. Lewis, K. Smith-Miles, and K. Phillips, “The school bus routing problem: An analysis and algorithm,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 10765 LNCS, no. 1, pp. 287–298, 2018.

[21] D. Cattaruzza, N. Absi, and D. Feillet, “Vehicle routing problems with multiple trips,” *Ann. Oper. Res.*, vol. 271, no. 1, pp. 127–159, 2018.

[22] W. Li, Y. Wu, P. N. R. Kumar, and K. Li, “Multi-trip vehicle routing problem with order release time,” *Eng. Optim.*, vol. 52, no. 8, pp. 1279–1294, 2019.

[23] A. Mingozzi, R. Roberti, and P. Toth, “An exact algorithm for the multitrip vehicle routing problem,” *INFORMS J. Comput.*, vol. 25, no. 2, pp. 193–207, 2013.

[24] F. Semiz and F. Polat, “Solving the area coverage problem with UAVs: A vehicle routing with time windows variation,” *Rob. Auton. Syst.*, vol. 126, p. 103435, 2020.

[25] J. Cordeau, H. Etudes, J. Desrosiers, and M. M. Solomon, “7. VRP with Time Windows,” in *The Vehicle Routing Problem*, 2000, pp. 157–193.

[26] M. Dorigo, G. Di Caro, and L. M. Gambardella, “Ant Algorithms for Discrete Optimization,” *Artif. Life*, vol. 5, no. 2, pp. 137–172, 1999.

[27] N. A. El-Sherbeny, “Vehicle routing with time windows: An overview of exact, heuristic and metaheuristic methods,” *J. King Saud Univ. - Sci.*, vol. 22, no. 3, pp. 123–131, Jul. 2010.

[28] A. Grasas, H. Ramalhinho, L. S. Pessoa, M. G. Resende, I. Caballé, and N. Barba, “On the

improvement of blood sample collection at clinical laboratories," *BMC Health Serv. Res.*, vol. 14, no. 1, 2014.

[29] A. Expósito, J. Brito, J. A. Moreno, and C. Expósito-Izquierdo, "Quality of service objectives for vehicle routing problem with time windows," *Appl. Soft Comput. J.*, vol. 84, Nov. 2019.

[30] A. H. Land and A. G. Doig, "An Automatic Method of Solving Discrete Programming Problems," *Econometrica*, vol. 28, no. 3, p. 497, 1960.

[31] D. R. Morrison, S. H. Jacobson, J. J. Sauppe, and E. C. Sewell, "Branch-and-bound algorithms: A survey of recent advances in searching, branching, and pruning," *Discret. Optim.*, vol. 19, pp. 79–102, 2016.

[32] M. Blocho, "Exact algorithms for solving rich vehicle routing problems," in *Smart Delivery Systems*, Elsevier, 2020, pp. 93–99.

[33] G. Laporte, "What you should know about the vehicle routing problem," *Nav. Res. Logist.*, vol. 54, no. 8, pp. 811–819, Dec. 2007.

[34] G. Clarke and J. W. Wright, "Scheduling of Vehicles from a Central Depot to a Number of Delivery Points," *Oper. Res.*, vol. 12, no. 4, pp. 568–581, Aug. 1964.

[35] M. Blocho, "Heuristics, metaheuristics, and hyperheuristics for rich vehicle routing problems," in *Smart Delivery Systems*, Elsevier, 2020, pp. 101–156.

[36] J. F. Cordeau, G. Laporte, M. W. P. Savelsbergh, and D. Vigo, "Chapter 6 Vehicle Routing," in *Handbooks in Operations Research and Management Science*, vol. 14, no. C, 2007, pp. 367–428.

[37] S. Lin, "Computer Solutions of the Traveling Salesman Problem," *Bell Syst. Tech. J.*, vol. 44, no. 10, pp. 2245–2269, 1965.

[38] P. M. Thompson and H. N. Psaraftis, "Cyclic transfer algorithms for multivehicle routing and scheduling problems," *Oper. Res.*, vol. 41, no. 5, pp. 935–946, 1993.

[39] A. Van Breedam, "An Analysis of the Behavior of Heuristics for the Vehicle Routing Problem for a selection of problems," 1994.

[40] G. A. P. Kindervater and M. W. P. Savelsbergh, "10. Vehicle routing: handling edge exchanges," in *Local Search in Combinatorial Optimization*, 2003, pp. 337–360.

[41] G. Laporte, "Classical Heuristics for the Capacitated VRP," in *The V*, 2002, pp. 109–128.

[42] F. Glover, "Future paths for integer programming and links to artificial intelligence," *C. 0~s. Res*, vol. 13, no. 5, pp. 533–549, 1986.

[43] R. Martí, J. A. Lozano, A. Mendiburu, and L. Hernando, “Multi-start methods,” in *Handbook of Heuristics*, vol. 57, 2018, pp. 155–175.

[44] F. Oudouar, M. Lazaar, and Z. El Miloud, “A novel approach based on heuristics and a neural network to solve a capacitated location routing problem,” *Simul. Model. Pract. Theory*, vol. 100, Apr. 2020.

[45] N. Yoshiike and Y. Takefuji, “Solving vehicle routing problems by maximum neuron model,” *Adv. Eng. Informatics*, vol. 16, no. 2, pp. 99–105, 2002.

[46] A. M. Anaya-arenas, T. Chabot, J. Renaud, and A. Ruiz, “Biomedical sample transportation in the province of Quebec: a case study,” *Int. J. Prod. Res.*, vol. 54, no. 2, pp. 602–615, 2015.

[47] J. Yi, “Vehicle Routing with Time Windows and Time-Dependent Rewards: A Problem from the American Red Cross,” *Manuf. Serv. Oper. Manag.*, vol. 5, no. 1, pp. 74–77, 2003.

[48] K. F. Doerner, M. Gronalt, R. F. Hartl, G. Kiechle, and M. Reimann, “Exact and heuristic algorithms for the vehicle routing problem with multiple interdependent time windows,” *Comput. Oper. Res.*, vol. 35, no. 9, pp. 3034–3048, 2008.

[49] P. Ghandforoush and T. K. Sen, “A DSS to manage platelet production supply chain for regional blood centers,” *Decis. Support Syst.*, vol. 50, no. 1, pp. 32–42, 2010.

[50] A. Mobasher, A. Ekici, and O. Ö. Özener, “Coordinating collection and appointment scheduling operations at the blood donation sites,” *Comput. Ind. Eng.*, vol. 87, pp. 260–266, 2015.

[51] F. G. Şahinyazan, B. Y. Kara, and M. R. Taner, “Selective vehicle routing for a mobile blood donation system,” *Eur. J. Oper. Res.*, vol. 245, no. 1, pp. 22–34, Aug. 2015.

[52] A. M. Anaya-Arenas, T. Chabot, J. Renaud, and A. Ruiz, “Biomedical sample transportation in the province of Quebec: A case study,” *Int. J. Prod. Res.*, vol. 54, no. 2, pp. 602–615, 2016.

[53] M. Toschi, E. Lanzarone, A. M. Anaya-Arenas, V. Bélanger, V. Nicoletta, and A. Ruiz, “A Fix-and-Optimize Variable Neighborhood Search for the Biomedical Sample Transportation Problem,” *IFAC-PapersOnLine*, vol. 51, no. 11, pp. 992–997, 2018.

[54] A. M. Anaya-Arenas, C. Prodhon, J. Renaud, and A. Ruiz, “An iterated local search for the biomedical sample transportation problem with multiple and interdependent pickups,” *J. Oper. Res. Soc.*, pp. 1–16, 2019.

[55] Z. Naji-Azimi, M. Salari, J. Renaud, and A. Ruiz, “A practical vehicle routing problem with desynchronized arrivals to depot,” *Eur. J. Oper. Res.*, vol. 255, no. 1, pp. 58–67, 2016.

[56] L. C. Coelho and G. Laporte, “Improved solutions for inventory-routing problems through valid

inequalities and input ordering," *Int. J. Prod. Econ.*, vol. 155, pp. 391–397, 2014.

[57] R. Martí, "Multi-Start Methods," in *Handbook of Metaheuristics*, 2006, pp. 355–368.

[58] A. M. Campbell and M. Savelsbergh, "Efficient insertion heuristics for vehicle routing and scheduling problems," *Transp. Sci.*, vol. 38, no. 3, pp. 369–378, 2004.

[59] "WSmart Route." [Online]. Available: http://wsmartroute.tecnico.ulisboa.pt/. [Accessed: 22-Dec-2020].

Annexes

A. Current routes

A1. Region A

Table A1.1 - Current solution of Regions A6 and A7.

Routes	Time taken (min)	Total Distance (km)	Departing time	Arrival Time
2-49-50-51-52-53-54-55-56-57-1	517	533	09h00	17h36
3-27-28-29-30-31-32-33-34-35-36-37-38-1	290	212	10h00	14h50
4-8-9-10-11-12-13-14-15-1	286	258	11h00	15h46
5-16-17-18-19-20-21-22-23-24-25-26-1	250	175	11h00	15h10
6-39-40-41-42-43-44-45-46-47-48-1	294	242	10h00	14h53
7-58-59-60-61-62-63-64-65-66-67-1	376	345	11h00	17h16
All routes	2013	1765	-	-

Table A1.2 - Current solution of the morning routes of Regions A1, A2, A3, A4 and A5.

Routes	Time taken (min)	Total Distance (km)	Departing time	Arrival Time
2-16-17-18-19-20-21-22-23-24-25-26-27-28-29-30-31-32-1	230	60	10h30	14h20
3-33-34-35-36-37-38-39-40-41-42-43-44-45-46-47-1	260	110	09h45	14h05
4-48-49-50-51-52-53-54-55-56-57-58-59-60-61-62-63-64-65-66-1	348	158	10h00	15h48
5-67-68-69-70-71-72-73-74-75-76-77-78-79-80-81-82-83-84-85-86-87-1	332	122	10h30	16h02
6-88-89-90-91-92-93-94-95-96-97-98-99-100-101-102-103-1	195	35	11h00	14h15
7-104-105-106-107-108-109-110-111-112-113-114-115-116-117-118-119-120-1	231	61	09h15	14h15
8-121-122-123-124-125-126-127-128-129-130-131-132-133-134-135-136-1	207	47	11h00	14h27
9-137-138-139-140-141-142-143-144-145-146-147-148-149-150-151-152-153-154-155-156-157-158-159-1	328	98	10h30	15h58
10-160-161-162-163-164-165-166-167-168-169-170-171-1	167	47	11h30	14h17
11-172-173-174-175-176-177-178-179-180-181-182-183-184-185-186-187-188-190-191-192-193-194-1	355	125	10h30	16h25
12-195-196-197-198-199-200-201-202-203-204-205-206-207-208-209-210-211-1	251	81	11h30	15h41
13-212-213-214-215-216-217-218-219-220-221-222-223-224-225-226-227-228-229-1	249	69	11h30	15h39
14-230-231-232-233-234-235-236-237-238-239-240-241-242-243-244-245-246-247-248-1	229	39	10h30	14h19
15-249-250-251-252-253-254-255-256-257-258-259-260-261-262-263-264-265-1	275	105	09h45	14h20
All routes	3657	1157	-	-

Table A1.3 - Current solution of the afternoon routes of Regions A1, A2, A3, A4 and A5.

Routes	Time taken (min)	Total Distance (km)	Departing time	Arrival Time
2-14-15-16-17-18-19-20-21-22-23-24-25-26-1	226	96	15h20	19h06
3-27-28-29-30-31-32-33-34-35-36-37-38-1	158	38	15h05	17h43
4-39-40-41-42-43-44-45-46-47-1	213	123	16h48	20h21
5-48-49-50-51-52-53-54-55-56-57-1	138	38	17h02	19h20
6-58-59-60-61-62-63-64-1	251	181	15h15	19h26
7-65-66-67-68-69-70-71-72-73-74-75-1	164	54	14h06	16h50
8-76-77-78-79-80-81-82-1	153	83	15h27	18h00
9-83-84-85-86-87-88-89-90-91-92-1	168	68	16h58	19h46
10-93-94-95-96-97-98-99-100-101-102-1	162	62	15h17	16h57
11-103-104-105-106-107-108-109-110-111-112-1	125	25	16h39	18h44
12-113-114-115-116-117-118-119-120-121-122-1	129	29	15h19	17h28
13-123-124-125-126-127-128-1	85	25	15h20	16h45
All routes	1972	822	-	-

A2. Region B

Table A2.1 - Current solution of Region B.

Routes	Time taken (min)	Total Distance (km)	Departing time	Arrival Time
2-5-6-7-8-9-10-11-12-1	541	502	10h00	19h01
3-13-14-15-16-17-18-19-20-1	424	355	11h00	18h04
4-21-22-23-24-25-26-27-28-29-30-31-1	642	590	10h00	20h42
All routes	1607	1447	-	-

B. Proposed routes

B1. Region A

Table B1.1 - Solution for Regions A6 and A7.

Routes	Time taken (min)	Total Distance (km)	Departing time	Arrival Time
2-10-9-38-64-65-67-49-1	485	536	09h48	17h53
3-22-21-20-8-26-13-14-11-12-58-1	277	246	09h24	14h01
4-41-27-40-28-44-43-15-25-24-23-19-18-17-1	240	170	09h43	13h43
5-16-29-45-57-46-37-33-32-34-36-35-54-55-30-42-1	388	335	09h37	16h05
6-56-53-51-52-50-48-47-39-1	361	371	09h33	15h34
7-31-66-63-62-61-60-59-1	323	334	09h41	15h04
All routes	2074	1843	-	-

Table B1.2 - Solution for the morning routes of Regions A1, A2, A3, A4 and A5.

Routes	Time taken (min)	Total Distance (km)	Departing time	Arrival Time
2-16-74-71-23-31-25-30-29-28-27-26-24-17-20-22-21-19-88-1	243	63	09h38	13h41
3-49-45-46-43-42-41-40-37-39-38-36-34-33-55-54-68-1	333	173	09h21	14h54
4-162-65-63-64-59-60-57-56-53-50-51-48-52-44-47-73-79-81-1	340	160	09h59	15h39
5-67-18-85-84-83-80-75-77-76-70-78-72-69-35-58-62-66-61-1	308	128	09h38	14h46
6-210-175-92-98-97-227-216-212-155-213-95-174-221-222-203-176-173-1	224	54	10h04	13h48
7-116-115-114-113-112-111-110-109-108-107-104-118-119-101-99-32-82-1	273	103	11h02	15h35
8-134-133-132-131-130-129-128-127-126-125-123-124-122-137-105-256-257-258-255-1	280	90	11h03	15h43
9-151-148-149-150-140-141-145-142-144-143-139-147-121-136-120-135-1	220	60	11h05	14h45
10-169-167-168-164-163-161-165-160-103-100-243-244-146-152-138-153-159-1	226	56	11h01	14h47
11-172-178-179-106-117-102-233-166-166-171-170-157-158-1	185	65	11h05	14h10
12-154-206-196-205-198-197-195-194-193-191-192-180-190-250-187-186-185-249-184-183-182-181-177-1	320	90	09h54	15h14
13-247-218-238-94-89-215-214-211-91-93-207-209-208-86-200-87-199-201-202-96-1	275	75	10h05	14h40
14-156-248-230-232-245-241-240-239-231-235-236-237-234-246-242-90-229-228-226-225-224-223-204-220-219-217-1	304	44	10h01	15h05
15-189-263-265-259-260-251-252-188-253-261-262-264-254-1	223	93	10h00	13h43
All routes	3754	1254	-	-

Table B1.3 - Solution for the afternoon routes of Regions A1, A2, A3, A4 and A5.

Routes	Time taken (min)	Total Distance (km)	Departing time	Arrival Time
2-16-22-20-21-17-14-24-23-1	146	66	14h43	17h09
3-39-40-59-58-15-25-1	231	171	15h58	19h49
4-28-91-84-85-80-77-78-79-76-60-61-102	245	125	15h47	19h52
5-72-73-37-38-35-36-31-34-33-122-46-47-1	214	94	14h54	18h28
6-125-45-44-93-42-43-68-70-1	117	37	16h37	18h34
7-57-54-55-56-52-53-51-48-49-82-81-1	172	62	15h46	18h38
8-126-69-65-63-64-62-41-111-112-1	171	81	15h48	18h39
9-29-30-75-67-66-97-1	97	37	15h16	16h53
10-32-27-92-83-87-90-89-86-88-74-71-1	174	64	16h19	19h13
11-110-96-95-115-94-99-101-1	99	29	15h41	17h20
12-114-113-127-128-100-104-108-109-119-98-120-121-1	157	37	16h12	18h49
13-50-26-106-103-123-124-118-117-116-105-107-19-18-1	188	58	14h49	17h57
All routes	2011	861	-	-

B2. Region B

Table B2.1 - Solution for Region B.

Routes	Time taken (min)	Total Distance (km)	Departing time	Arrival Time
2-9-8-7-6-5-1	416	383	11h43	18h39
3-10-15-16-19-18-17-14-13-11-12-21-1	487	397	10h02	18h09
4-20-23-24-25-27-26-29-30-31-22-1	542	465	09h56	18h58
All routes	1445	1245	-	-