Mathematical Modelling to Support Blood Collection for the Welsh Blood Service

Emily Paige Williams
School of Mathematics
Cardiff University

A thesis submitted for the degree of

Doctor of Philosophy

May 2022

Abstract

Human blood is a scarce resource and its role in healthcare is fundamental, with donated blood saving the lives of many on a daily basis. The blood supply chain is responsible for the transfer of blood from donor to the recipient, but the availability of such an invaluable resource as human blood is ultimately attributable to the many voluntary donors. Thus, the efficiency of the collection of donated blood is crucial to the downstream effectiveness of the blood supply chain.

Working in partnership with the Welsh Blood Service, our aim is to create a decision support tool to aid the scheduling process to match supply and demand of blood products, whilst minimising costs and wastage in the system. We present an integer linear programme model that consists of two stages. The first stage schedules mobile blood donation clinics, considering over 300 locations, with the objective to minimise both the number of clinics scheduled within the planning horizon and the amount of blood collected that exceeds the demand. The second stage assigns workers to each scheduled clinic, with the objective of minimising costs such as overtime costs. Both stages of this scheduling model are developed in Python and are solved using PuLP - an open source Python package which utilises COIN-OR CBC solver.

Test instances are designed and the experimental results are presented which demonstrate the effectiveness of the two-stage model to improve cost and time efficiencies of the collections process at the Welsh Blood Service, in addition to enabling the matching of supply to demand. Finally, some insights regarding the staffing levels of each region are discussed.

Acknowledgements

Firstly, I would like to thank my academic supervisors, Prof. Paul Harper and Prof. Daniel Gartner, for your unwavering support and guidance throughout the duration of this project. I will be forever grateful for your patience, understanding and encouragement.

I would like to extend my gratitude to all at the Welsh Blood Service who have so kindly given their time to support my research, and for always making me feel welcome. More specifically, thank you to Leanne Morgan, Lisa Davies and Samantha Silcox for your generosity and patience in helping me to understand the planning process at the WBS and for answering my many questions; thank you to Cath O'Brien for making this industrial partnership possible, and to Paul Wilkins for your invaluable insights. Additionally, I'd like to thank Sian James and Nick O'Sullivan for facilitating my project at the WBS. A very special thank you to Jayne Davey, my industrial supervisor, for your kindness, for championing my research and for your regular words of encouragement.

I am grateful to the School of Mathematics staff and the PhD cohort for upholding such a friendly and supportive environment, and to the members of WiM for helping to create a wonderful and uplifting community. Thank you to Dr. Vincent Knight, for introducing me to Python and Siren, and to Dr. Geraint Palmer, Dr. Nikoleta Glynatsi and Dr. Henry Wilde for sharing your Python knowledge.

Furthermore, I would like to thank KESS2 for their facilitation and funding of this project – without whom this research would not have been possible – with a special thank you to Esther Meadows for your advice and assistance over the course of the PhD.

On a personal note, I would to thank my friends and family for your support, patience and welcome distractions over the last four years. More specfically, thank you to my parents and Louise for always helping in any way you can, and thank you to Karina for your unrelenting optimism and generosity. Finally, a very special thank you to Dr. Emma Aspland, my biggest cheerleader – I truly could not have completed this PhD without your support and friendship.

Dissemination

Publications

- Modeling of the collections process in the blood supply chain: A literature review. E.P. Williams, P.R. Harper and D. Gartner. IISE Transactions on Healthcare Systems Engineering, 2020. [101]

Presentations

- Modelling to Support the Redesign of the Welsh Blood Service Supply Chain System (Poster). ORAHS, Oslo (Norway), August 2018.
- Modelling to Support the Redesign of the Welsh Blood Service Supply Chain System (Poster). OR60, Lancaster (United Kingdom), September 2018.
- Scheduling Blood Donation Clinics to Match Supply and Demand. SIAM Student Chapter National Conference. Manchester (United Kingdom), June 2019.
- Scheduling Blood Donation Clinics to Match Supply and Demand. EURO, Dublin (Ireland), June 2019.
- Scheduling Blood Donation Clinics to Match Supply and Demand. ORAHS, Karlsruhe (Germany), July 2019. *Winner of Steve Gallivan Award for 'Best Presentation by an Early Career Operational Researcher'.*

- Scheduling Blood Donation Clinics to Match Supply and Demand. EURO PhD Summer School; 'Operational Research for Value-based Health Care'. Lisbon (Portugal), September 2019.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

This chapter serves as an introduction to the background of this research project, conducted as the product of a research partnership with the Welsh Blood Service (WBS), funded by both the WBS and Knowledge Economy Skills Scholarships (KESS 2). The chapter begins with a history of blood transfusion described in Section 1.1, followed by a discussion about the demand for blood products in Section 1.2, and subsequently, the blood supply chain is described in Section 1.3. The Welsh Blood Service and how they collect blood donations are presented in Section 1.4 alongside details of the research partnership. Research aims of this project are discussed in Section 1.5, followed by a thesis outline in Section 1.6

1.1 History of Blood Transfusion

Human blood is a scarce resource and its role in modern healthcare is fundamental, with donated blood saving the lives of many on a daily basis. The first reported successful human to human blood transfusion was performed by James Blundell, a British obstetrician. There is debate about when this blood transfusion occurred, but it likely took place sometime between 1818 and 1829, with the transfusion for treatment of postpartum haemorrhage [31]. Prior to this, many experiments were conducted transfusing blood between animals, including between different species, and even from animals such as sheep to humans as early as 1667. Many of these experiments resulted in deaths [18]. However, James Blundell was aware of the research of John Leacock, who theorised that blood transfusions between the same species could be safe. Blundell undertook his own research, transfusing blood between dogs, before his first attempted blood transfusion on a human [31]. In total, Blundell was responsible for ten human to human blood transfusions, with five of these being successful. However, blood types and compatibility were not discovered until much later.

Human blood types, A, B and O, were discovered in 1901 by Austrian physician, Karl Landsteiner, along with the conclusion that when blood of different types were mixed, this often led to clotting. Landsteiner was awarded the Nobel Prize for his work in 1930 [3]. His colleagues Alfred Decastello and Adriano Sturli discovered the fourth blood type, AB, shortly after in 1902. In 1937, Karl Landsteiner and Alexander Weiner, with later contributing findings made by Philip Levine and Rufus Stetson in 1940, discovered the Rh blood group system (named in error after the Rhesus Monkey) [28]. The Rh blood groups consider the antigens present in human blood, and their names are given by the presence of antigen Rh(D) i.e. Rh positive indicates the presence of Rh(D), whilst Rh negative indicates the lack of antigen Rh(D). Combined with the ABO blood group system, these two blood group systems form the basis of blood type compatibility, as displayed in Table 1.1.

Advances were made in the understanding of anticoagulants, such as sodium citrate proposed by Richard Lewisohn [57], and their vital role in preserving blood products so that they could be stored for some days before being used. This facilitated blood banks being developed, with the first known blood bank opened in Leningrad Hospital in Russia in 1932. Many other blood banks were set up in the following years, with blood transfusions becoming increasingly common in the 1940s as a result of the demand caused by World War II [18]. Further progress is made

		Donor's Blood Type							
		O−	**O+**	**B−**	**B+**	**A−**	**A+**	**AB−**	**AB+**
Recipient's Blood Type	**O−**	✓	✓	✓	✓	✓	✓	✓	✓
	O+	✓	✗	✓	✗	✓	✗	✓	✗
	B−	✓	✓	✗	✗	✓	✓	✗	✗
	B+	✓	✗	✗	✗	✓	✗	✗	✗
	A−	✓	✓	✓	✓	✗	✗	✗	✗
	A+	✓	✗	✓	✗	✗	✗	✗	✗
	AB−	✓	✓	✗	✗	✗	✗	✗	✗
	AB+	✓	✗	✗	✗	✗	✗	✗	✗

Table 1.1: Blood Type Compatibility

through the mid to late twentieth century developing collection, storage and testing technologies [26], in addition to the separation of blood into its components which enables each donation to treat several patients.

1.2 Demand for Blood Products

The World Health Organization (WHO) estimates that 118.4 million blood donations are collected each year globally [67]. Regarding the demand, according to a study conducted by N. Roberts et al. [82], in 2017 the global blood need was approximately 304.7 million blood products whilst the global blood supply was estimated to be 272.3 million blood product units. The disparity between global supply and demand is clear from these figures, with shortages more likely to occur in low-income countries [67].

After collection, blood donations are usually separated into its components, including red blood cells, platelets, and plasma, with each of these products used for different treatments. Red blood cells can be used to treat all types of anaemia (including anaemia caused by rheumatoid arthritis or cancer), while red blood cell transfusions are frequently used to replace heavy blood loss as a result of childbirth, surgery or an accident [19]. Platelets are often utilised to treat bone marrow failure, leukaemia, and following a transplant or chemotherapy. Finally, plasma is commonly given to trauma and burn patients, in addition to being important in treatments for immune deficiencies. With the wide variety of life-saving treatments enabled by blood donation, it is a critical element of modern-day healthcare.

Most blood donations in the world today are voluntary; as of 2018, the WHO estimates that 79 countries collect over 90% of their blood supply via unpaid donors [67]. However, there were still 56 countries that receive more than 50% of their blood supply from either paid donors or from the family of the patient.

1.3 The Blood Supply Chain

The blood supply chain is responsible for the process of transportation of blood products from the vein of a donor to the vein of a patient. The success of the supply chain is of utmost importance, with inefficiencies and errors potentially leading to fatalities. The structure of the blood supply chain model may vary depending on location, politics, costs, etc. but the ultimate goal is universal: to satisfy the demand at minimal cost and minimal wastage [70].

The blood supply chain is commonly divided into the following four echelons (Osorio et al. [68]):

- **Collection:** This is the beginning of the blood supply chain and involves either fixed or mobile blood donation clinics, or a combination of both. Blood products may be collected in the form of either an individual blood component via a process called apheresis, or whole blood which is the most common form of donation. From a donor's perspective, a clinic typically follows the process displayed in Figure 1.1, with queues often arising between the activities (which we will refer to as workstations). Eligibility screening is necessary as there are many restrictions on who may donate blood [92]. Screening often becomes a prolonged process with many health-related questions and a test for iron levels within the donor's blood. The collection echelon ends with all blood donation units transported to a blood processing centre.

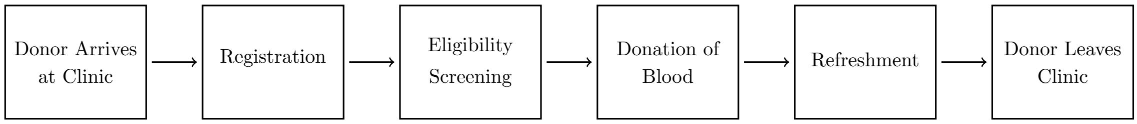


Figure 1.1: Typical Path of a Donor at a Donation Clinic

- **Production:** This takes place at a blood processing centre, where donated units of blood are tested and separated into various components (red blood cells, platelets and plasma) as required. The production of platelets depends on the amount of time since the donation, as platelets must be separated from whole blood shortly after collection. This echelon ends with blood being packaged ready for distribution and moved to storage.

- **Inventory:** Storage of blood products may either take place at a blood processing centre or a stock-holding unit. Each type of blood product has unique shelf-lives and specific storage requirements, with platelets being the most complicated; platelets must be kept in an agitated state at ambient temperature for a maximum of approximately 5 days.

- **Distribution:** This echelon consists of the preparation of orders of blood products and the transportation of such orders to the respective hospitals. Decisions involved include the dates and blood types of dispatched products, due to blood compatibility and possible limited inventory.

1.4 The Welsh Blood Service

The research presented in this thesis was undertaken as part of a project funded by both Knowledge Economy Skills Scholarships (KESS 2) and the Welsh Blood Service (WBS). KESS 2 facilitates collaborative projects between organisations and academia to support the development of key technologies in Wales, funded by the European Social Fund via the Welsh European Funcding Office (WEFO). The WBS is a division of the Velindre University NHS Trust and is responsible for the blood supply chain in Wales, UK. The service previously only covered South West and

South East Wales but as of 2016, it is an all-Wales blood service. The WBS collects blood from the general public on a voluntary, non-remunerated basis, and processes, tests, and distributes blood products to hospitals across Wales.

The Welsh Blood Service has a total of 577 employees as of September 2021, with approximately 125 of these employed as whole blood donation clinic staff. The organisation operates mainly mobile blood donation clinics, with only one fixed clinic located at the headquarters of the organisation. These mobile clinics consist of both ‘community’ clinics (set in a local venue, often at a monetary cost) and ‘trailer’ clinics (set in a mobile trailer, often parked at a supermarket or business). The vast majority of these clinics operate for one day only each time the clinic is scheduled, whilst some may operate over several consecutive days each time - we refer to these as multi-day clinics. The current collection model also involves clinic ‘tours’ which consist of more rural clinics being operated over consecutive days. Due to the location of these clinics, the clinic staff stay overnight at accommodation local to the clinics to minimise travel time. The WBS also operate several apheresis clinics each week for donation of platelets at a clinic based at the organisation’s headquarters.

1.4.1 Facility and Clinic Locations

The WBS operate mobile whole blood clinics in around 350 different locations across Wales. Each clinic is assigned to a base (usually the closest in geographical distance) i.e. where equipment is stored and where staff are typically based. There are a total of four bases across Wales, located in Bangor, Wrexham, Dafen and Talbot Green (WBS Headquarters), resulting in four regions. Figure 1.2 displays the location of the four bases (represented by black pinpoints) with each clinic location represented by a coloured circle, the colour indicative of the base of which the clinic is assigned i.e. blue represents clinics assigned to Bangor base, yellow represent the clinics assigned to Wrexham base, red to Dafen base and green to Talbot Green (HQ).

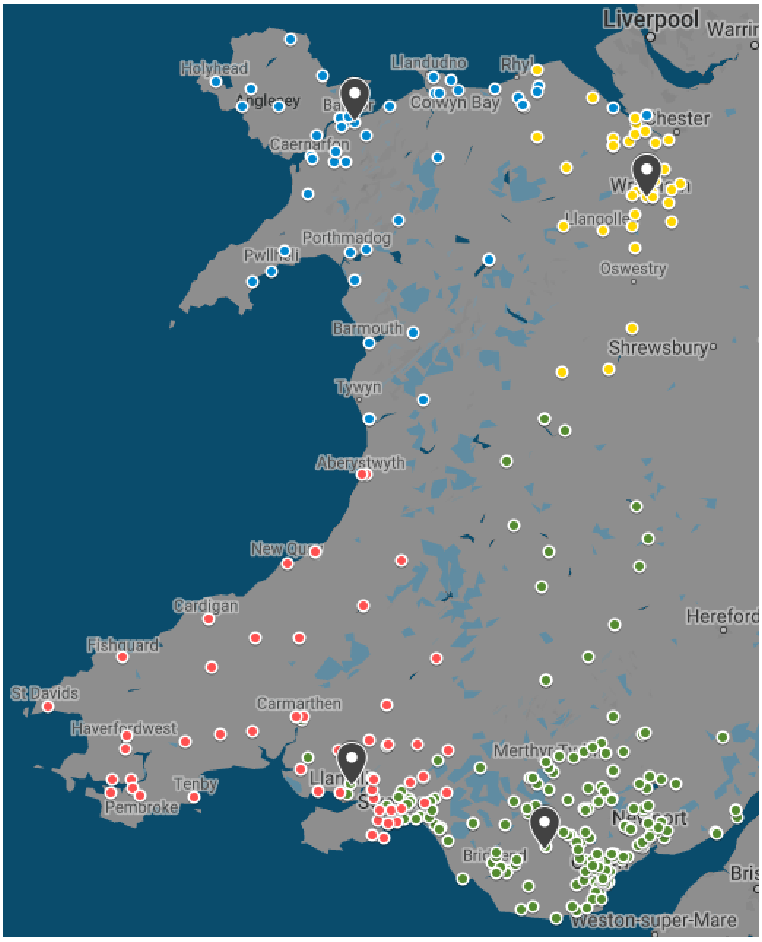


Figure 1.2: Welsh Blood Service Clinics by Region

All of the blood products are processed for production and tested at the headquarters in Talbot Green, requiring all collected blood units to be transported to HQ at the end of each working day. The storage of processed blood products is mostly at HQ also, with some stock transported to a stock-holding unit in North Wales for faster delivery to hospitals located in the north of Wales, as and when required.

1.4.2 Frequency of Clinics

Whole blood donation clinics have a variety of regulations to ensure a safe service is provided. One of these regulations is a temporal restriction between consecutive donations: For male[1] donors this means that there must be a minimum of 12 weeks between consecutive donations, whereas there is a minimum of 16 weeks for female

[1]Where the terms male and female are used, we are referring to those assigned this sex at birth.

donors. Due to this temporal restriction, there are also restrictions on the frequency of mobile clinics, with most scheduled no more frequently than once every 16 weeks. Some clinics may have a significantly large donor panel and thus may operate several times within a 16-week period, with a different section of its donor panel targeted each time. The operating hours for each mobile clinic are predetermined, with time for travelling from team base, set-up and set-down of the clinic taken into consideration. The majority of clinics run for one day, whilst the remaining clinics may operate over several consecutive days each time. This is usually the case in clinics that have a large donor panel and/or require several hours in travelling to and from the clinic for staff, with the latter sometimes resulting in staff requiring overnight accommodation.

1.4.3 Collections Planning

Currently, the clinic planning process is tedious and time-consuming, and carried out in four-week planning periods, relying on various sources of data and local knowledge to form each plan. Each clinic is given a collection estimate based on an average of the previous three days that the clinic was operated, and this is taken into account during the scheduling process. However, fulfilling staff contractual hours is also a goal in addition to meeting the blood product demand i.e. ensuring a minimum number of clinics are scheduled in each region per four-week period. This can result in more clinics being scheduled than is necessary to reach demand, potentially causing overcollection of blood or the additional associated costs of operating more (and often smaller, less efficient) clinics to meet the demand.

Similarly to clinic schedules, workforce schedules are currently completed manually and in four-week periods. Each clinic has various skill mix requirements, depending on the capacity of the clinic and venue type, and these must be considered when workforce schedules are created. Training of clinic staff must also be considered to ensure that all required certification is valid.

1.4.4 Research Partnership

The research presented in this thesis was conducted as a project funded by both Knowledge Economy Skills Scholarships (KESS 2) and the Welsh Blood Service. The Welsh Blood Service has provided a project supervisor throughout the duration of the project to oversee work and ensure all research aligns with the requirements of the organisation. Through discussion with the WBS, it was established that the organisation sought to improve clinic and workforce planning processes to reduce costs and wastage in the system, such as resources' time and wastage of blood products from overcollection.

Despite the success of the blood supply chain ultimately being dependent on the efficiency of the collection of donated blood, there is a current lack of existing research focussing on optimisation of the whole blood collections process (as shown in Chapter 2). This gap in the literature overlaps with the aims of the WBS. Therefore, throughout this research we focus exclusively on the collection echelon, more specifically the collection of whole blood donations.

More detailed descriptions of the Welsh Blood Service collection model are presented in further chapters, namely Chapter 3 and Chapter 4.

1.4.5 Available Data

The Welsh Blood Service provided data to aid research where possible, with NHS data protection compliance so that donors cannot be identified. All clinic data was made accessible such as clinic names, locations, venue capacities, and resource requirements. Due to the anonymous nature of historical bleed figures, all figures from 2017 to 2020 were provided; these include how many units of blood were collected from each clinic and how many donors attended. Similarly, all blood issuing data was also provided by the WBS for the period from January 2017 to March 2020, which includes the age of each blood product at the time of issue.

The WBS also provided an anonymised dataset of all donation clinic staff, with identifiable information such as names and addresses removed. This data includes staff roles, skill level, agreed working day patterns and contracted hours. While only salary bands for each employee were provided by the WBS and not a specific pay rate, information regarding salary band pay rates and annual leave entitlement is openly accessible online, published by the NHS.

1.5 Research Aims

Having introduced the partner organisation and their requirements, we now present our research aims. The research detailed in this thesis aims to answer the following research questions:

1. How can mathematical modelling help to schedule the Welsh Blood Service's blood donation clinics more efficiently?

2. How can mathematical modelling help to schedule the clinic-based workforce at the Welsh Blood Service?

3. Can these mathematical methods be integrated into a decision support tool for planners at the Welsh Blood Service to use?

These research questions were identified through in-depth discussion with the Welsh Blood Service about their needs and the strategic direction for the organisation. Research question one arose from the goal of the WBS to minimise wastage of expired blood products, usually caused by overcollection of blood donations. Research question two follows a similar basis to research question one, with the WBS requiring staffing-related costs to be reduced. Finally, research question three stems from the desire of the WBS to have a usable tool to improve clinic and workforce scheduling.

1.6 Thesis Outline

This thesis consists of seven chapters, together attempting to answer the research questions described in Section 1.5. This chapter has introduced the Welsh Blood Service and a general overview of how they collect blood from donors across Wales, in addition to a summary of the blood supply chain and the history and importance of blood transfusion. Chapter 1 has also provided the research questions of this thesis.

Chapter 2 is a review of all relevant literature to the optimisation of the collection echelon of the blood supply chain. It provides a detailed taxonomy of any methods utilised, objectives of any research described in the publications, and an outline of the planning decision levels considered. This chapter identifies gaps in the literature and therefore areas of research that require further exploration, namely optimisation of workforce planning.

Chapter 3 explains the current clinic scheduling practice of the Welsh Blood Service, detailing aspects such as the venues, frequency and availability of clinics. The blood supply and demand are also discussed. This is followed by the limitations of the current practice, and a mathematical formulation of the Blood Donation Clinic Scheduling Problem in the form of a linear programme to optimally schedule clinics over a given planning horizon. Three alternative objective functions are presented along with many constraints to ensure the output clinic schedule is as realistic as possible. This formulation addresses research question one.

Chapter 4 details the current practice at the Welsh Blood Service of workforce scheduling for clinic-based staff. All clinic roles and related skill requirements are introduced, in addition to how overtime is managed and utilised at the Welsh Blood Service. Issues caused by the current workforce scheduling practice at the WBS are discussed before the mathematical formulation of the Blood Collection Workforce Scheduling Problem is presented. This problem is also formulated as a linear programme to optimally assign workers to a clinic schedule, with two alternative objective functions introduced and followed by various constraints to ensure that the workforce schedule solution is feasible for the WBS. This formulation addresses research question two.

Chapter 5 presents a prototype clinic scheduling model developed in Microsoft Excel using the OpenSolver add-in. The aims of this prototype are discussed in addition to the limitations of the model, and how this model addresses research question three. This chapter also presents the development of the formulations introduced in Chapter 3 (the Blood Donation Clinic Scheduling Model) and Chapter 4 (the Blood Collection Workforce Scheduling Model) as mathematical models in Python, utilising the open source linear programme library PuLP and solved using COIN-OR. The design of various test instances are introduced for both of the models.

Chapter 6 presents the experimental results for the test instances described in Chapter 5, with both computational and solution-related results of the BDCSM and the BCWSM. Key findings are discussed alongside and how these models improve upon current practice at the WBS, addressing research questions one and two.

Chapter 7 summarises the work of all previous chapters and provides answers to the research questions introduced in Section 1.5. The discussion section of this chapter addresses limitations related to implementation of the model. Following this, identified areas related to optimisation of clinic and workforce scheduling that would benefit from further work are explored.

Chapter 2

Literature Review

This chapter discusses literature that considers modelling of the collections process of the blood supply chain. Currently, the vast majority of literature concerns the inventory stage (Osorio et al. [68]) with little in-depth research carried out on the collections process. Therefore, the aim of this chapter is to evaluate the existing literature that deals with modelling of blood collection, provide a detailed classification of the selected articles, and thus identify any areas that may benefit from further research.

The remainder of this chapter is structured as follows: in Section 2.1, we describe how we conducted the structured search and provide an overview of previous literature reviews. In Section 2.2, we describe the relevant characteristics of blood collections. We demonstrate how the retrieved articles from Section 2.1 were classified into the various categories. Section 2.3 closes our literature review with conclusions.

2.1 Selection Criteria and Previous Reviews

2.1.1 Selection Criteria and Search for Relevant Literature

We search for journal publications from the Clarivate Analytics Journal Citation Report (JCR) in the subject categories of Health Policy and Services (HPS), Medical Informatics (MI), Industrial Engineering (IE), as well as Operations Research and the Management Sciences (OR/MS). These categories are selected due to the complex nature of the blood supply chain, with modelling of such arising from various fields and disciplines. The rationale for choosing these categories is to capture literature from a range of perspectives, whether it be policy decisions and service improvement (HPS), information systems and data mining (MI) or quantitative healthcare engineering (IE). The inclusion of these categories, together with OR/MS, yields a thorough analysis of research contributing to improvement of the collection echelon of the blood supply chain.

Using a structured search string, Scopus provided a base set of articles with the search mainly focusing on “blood collection” and “blood donation”. The search excludes all publications before 1996 and produced a total of 125 results in October 2021. All articles that did not specifically refer to the collections process were considered irrelevant, as were all articles that did not contain quantitative analysis or discuss optimisation of the blood supply chain. A forwards and backwards search from relevant articles was then conducted (as suggested by Webster & Watson [100]) and all relevant results collated. In carrying out our forwards search, we not only retrieved journal publications citing the articles from the original search but also a PhD thesis (van Brummelen [98]) which we decided to include in our final (relevant) set of publications. Collectively, this gives a total of 68 articles, as shown in Figure 2.1. This is an expanded literature search from the search we conducted in February 2019 [101] using the same search string, as displayed in Appendix A. Since February 2019, there has been an additional 22 relevant publications: 12 of these were results from the search string, six results were from the forwards search, and four results from the backwards search[1]

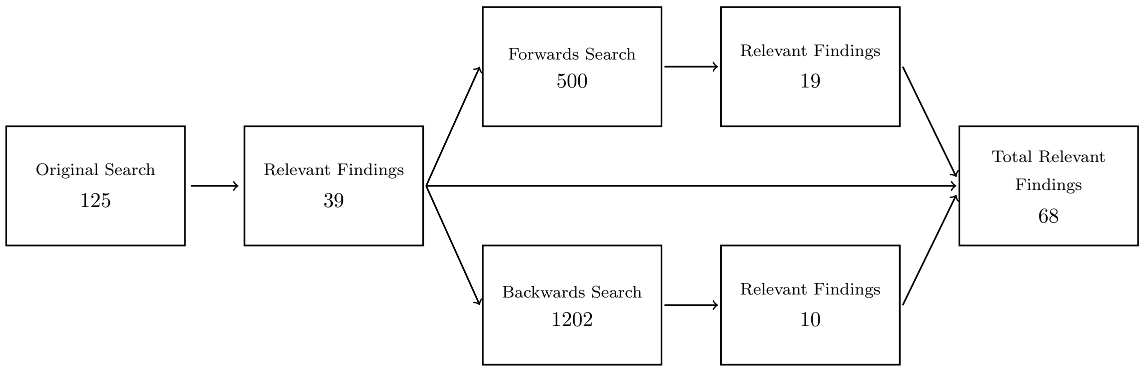


Figure 2.1: Scopus Search Results

2.1.2 Previous Reviews

There is a vast amount of literature that addresses modelling of the blood supply chain, however, only a minority of articles focus on the collection echelon. From our findings, there are five relevant reviews that address modelling of blood collection from donors; [11, 12, 14, 68, 75]. The most specific of these is written by Baş Güre et al. [11] who state that there are still several aspects of blood collection that have not yet been explored from an optimisation perspective, but specifically focus on donation appointment scheduling.

Baş Güre et al. also published another review [12], prior to the above article, which studies existing literature surrounding the blood supply chain in general, categorising by echelon and perspective of research. Similarly, Osorio et al. [68] provide a review of the whole blood supply chain, with a brief insight into the literature considering the collections process, categorisation of its planning decisions, and quantitative models of the process. In both reviews, the section addressing collections however is significantly smaller than those regarding other echelons of the supply chain, indicating that there is less existing research covering this particular

[1] Please note that Figure 2 in our published literature review [101] contained an error which has been corrected in this thesis; a publication was incorrectly included as a result from the original search. The relevant findings from the original search were in fact 27 (instead of 28) and the relevant findings from the forwards search were 13 (instead of 12).

echelon. However, Beliën and Forcé [14] deliver a review of the whole blood supply chain also, but instead categorise by methods used, blood product considered, etc. rather than by echelon of the supply chain. This makes it difficult to identify areas within sections of the blood supply chain that require further research.

Lastly, Pirában et al. provide a more detailed (and more recent) taxonomy of all literature involving models of the blood supply chain [75], with a breakdown of research by various categories such as methods, echelon(s), blood products considered, and objectives. This review identifies gaps in the existing literature such as the lack of studies that consider 'the design and application of collection policies to prevent inventory levels from exceeding the demand'.

The rest of this chapter provides a detailed taxonomy of articles, specifically focussing on the collection echelon of the blood supply chain. It includes all existing research on blood collection that consider an OR approach, including those from an interdisciplinary perspective, and provides an up-to-date analysis of surrounding literature due to the significant increase in publications within this field over recent years, as demonstrated by Table 2.1. Thus, our literature review presents an in-depth discussion of current literature on the collections process and clearly identifies areas that require further research.

2.2 Classification of Literature

Employing the selection criteria, a total number of 68 articles were retrieved and considered relevant for this review. The articles can be categorized by publication year as given in Table 2.1. The table reveals that this field of research is becoming increasingly popular, since 84% of the relevant articles were published within the last seven years. Earlier publications mainly consist of more simplistic statistical analyses, such as the impact of sharing blood donor deferral registries to streamline donation clinics [29]. However, in 1996 Jacobs et al. [50] published an integer programming method to advise the American Red Cross on whether to relocate a permanent facility. In more recent years, integer programming methods remain popular in optimising blood collection, in addition to many other OR methods including simulation and machine learning. The dramatic increase in research in optimisation of blood collection may be attributed to the technological advancements which enable the growing range of OR methods to be utilised effectively.

Table 2.1: Number of Articles by Publication Year

Time Period	Before 2005	2005–2007	2008–2010	2011–2013	2014–2016	2017–2019	2020–Present	Total
Number of Articles	2	2	2	5	14	27	16	68

The articles can also be classified by geographic location, as seen in Table 2.2. Here, articles are classified under either the location of any mentioned case studies, or that of the first named author. The majority of research in blood collections has been based in Asia, and most of these are in developing and emerging nations (such as Iran) and consider the blood collections process in the event of disasters (from natural or man-made causes) e.g. earthquakes. However, there is a lack of research regarding the blood supply chain in the event of disasters from elsewhere in the world, and this suggests that there is further research to be done to aid the blood services in countries who may face similar circumstances. There are a significant amount of articles concerning modelling of blood collection from Europe and America (North and South America collectively) which indicates that this is indeed a worldwide issue.

Table 2.2: Number of Articles by Continent

Continent	America	Asia	Europe	Rest of World	Total
Number of Articles	11	39	17	1	68

The articles may also be classified according to their respective JCR (Journal Citation Report) category to illustrate the approach and perspective of the research. The four subject categories are Health Policy and Services (HPS), Industrial Engineering (IE), Medical Informatics (MI) and Operations Research and Management

Sciences (OR/MS). A total of 47 articles are listed in Table 2.3, with the remaining 20 articles from other journals outside of these categories (all of which are from the forwards and backwards search) reinstating the multi-disciplinary nature of blood collection. The OR/MS category is clearly the most popular of the four categories; this conveys the usefulness of OR methods to tackle modelling of blood collection from donors.

Table 2.3: Articles by JCR Category

JCR Category	Articles	Total
HPS	[6, 29, 71, 103]	4
IE	[11, 33, 61, 63, 68, 73, 84, 90]	8
MI	[27, 96, 102]	3
OR/MS	[9, 10, 13, 14, 16, 22, 25, 34, 35, 41, 43, 44, 46, 50–52, 58, 60, 65, 69, 72, 75, 76, 78, 79, 81, 85, 86, 88, 89, 91, 104]	32

In the next subsections, a classification framework for these articles will be provided, categorising into functional areas, methods and approaches, planning decision level, and whether a case study was incorporated.

2.2.1 Functional Areas Considered

We now classify the articles by the functional area in which the research aims to improve/target. The functional areas have been organised into the following categories:

- **Appointment Scheduling:** All articles that discuss an appointment scheduling policy, framework, or optimisation of appointments.
- **Collection Policy:** This covers all research and analysis of the way in which blood is collected from donors, including eligibility/deferral of donors and also the collection strategy.
- **Crisis Situation:** This involves all research into the blood supply chain from the perspective of a crisis occurring and emergency aid being required e.g. natural disasters such as earthquakes.

- **Donor Demographics:** This includes all analysis and research of donor demographics such as donor behaviour, location, age, blood type, etc, and donor motivation.
- **Location/Clinic Planning:** This category includes all research which considers the location of either temporary or permanent facilities used for blood collection (mostly location of donation clinics). Both the allocation of clinics and relocation of facilities are categorised under this.
- **Staff Utilisation:** This category includes all articles which consider the allocation of staff to clinics and also analysis of staff level requirements and skill mix.
- **Vehicle Routing:** All articles that consider routing of vehicles that transport blood and resources for blood collection.

Note that articles which do not fit into any of the above categories are literature reviews and thus not proposing any particular methods or action.

Table 2.4: Functional Areas of Research

Functional Area	**Articles**
Appointment Scheduling	[6, 13, 63, 98, 103]
Collection Policy	[6, 9, 10, 27, 29, 39, 44, 60–62, 69, 71, 72, 87, 88, 102]
Crisis Situation	[25, 33–35, 39, 42, 51, 52, 59, 76, 78, 81, 86–91, 93, 95]
Donor Demographics	[5, 6, 27, 35, 43, 46, 56, 65, 79, 96]
Location/Clinic Planning	[5, 8, 9, 22–25, 30, 33–35, 39, 41, 42, 44, 46–48, 50–52, 61, 70, 76–79, 81, 84–86, 88–91, 93, 95, 104, 105]
Staff Utilisation	[5, 6, 17, 30, 58, 96, 98]
Vehicle Routing	[41, 59, 63, 73, 77, 81, 85]

2.2.1.1 Appointment Scheduling

Appointment Scheduling is the least popular category, as revealed by Table 2.4. While many donation sites worldwide accommodate for unbooked or "walk-in" donors, appointment systems provide an opportunity to control the arrival patterns of donors and better manage resources and inventory. Mobasher et al. [63]

consider how many donations should be collected within certain time intervals in accordance with vehicle routing to maximise donations viable for platelet production. Alfonso et al. [6] place more of an emphasis on appointment strategies regarding frequency of apheresis donations throughout the day, whilst Baş Güre et al. [13] focus on pre-allocating appointment slots to each blood type at a permanent facility. Van Brummelen [98] considers the combination of appointments and "walk-ins" and produces an optimal appointment schedule based on minimising waiting times for donors. Yalçındağ et al. [103] present a risk-averse framework for appointment scheduling under uncertain donor arrivals.

More generally, the literature surrounding appointment scheduling in healthcare neglects mainly strategic decisions but also most tactical decisions; according to Ahmadi-Javid et al. [4] the vast majority of publications concerning optimisation of outpatient appointments focus on an operational level. Contrastingly, the publications focusing on blood collection take a more tactical approach. Ahmadi-Javid et al. [4] also discuss the increasing interest in appointment scheduling, as over 73% of the publications considered in their review were produced between 2012 and 2016 (the most recent at the time of publication). Despite this, there is still a significant lack of research applied to blood donor appointment scheduling.

There is also a lack of research that considers combining various aspects of appointment scheduling. Reducing waiting times, matching supply to demand (including blood type specific demand) and managing donor arrivals are all key elements of effective and efficient blood collection, yet research within appointment scheduling often focus on just one of these goals.

2.2.1.2 Collection Policy

Collection policy is considerably widely researched, with a significant amount of these articles considering how many units of blood should be collected at a clinic [9,39,44,60–62,69,71]. This is considered from a variety of perspectives, such as the number of mobile clinics to deploy [44], how much to collect at each clinic under various scenarios [9, 87], or when to stop collecting each day [62]. Lowalekar & Ravi [61] use the Theory of Constraints (TOC) thinking process to evaluate the collections process associated with a blood bank in Chennai, India, to identify areas for improvement in the collection policy and thus improve inventory management. The TOC thinking process is a 'process improvement methodology that emphasizes the importance of identifying the "system constraint" or bottleneck' [97].

Donor deferrals are observed [6, 27, 29, 61], with Custer et al. [27] evaluating blood safety and policy decisions to assess the impact of deferrals on inventory. Xu et al. [102] consider the optimal interval between consecutive blood donations from male donors, ranging from eight weeks to 12 weeks. Various methods of blood product collection are also studied [6,69,71,72,88], with Osorio et al. [69] optimising the balance between whole blood and apheresis, considering cost and the number of donors required to reach demand. Ayer et al. [10] discuss the manufacturing of a specific blood product (cryoprecipitate) and how best to collect blood donations for this.

2.2.1.3 Crisis Situation

Of the 20 articles that address the blood supply chain in the event of a crisis situation, the majority of these also include location/clinic planning [25, 33–35, 39, 42, 51, 52, 76, 78, 81, 86, 88–91, 93, 95] with other articles considering alternative areas such as vehicle routing [59] or collection strategy [81, 87]. This is due to the importance of facilities (collection and processing) being located in an accessible area for a responsive and reliable supply chain. The aims of these articles vary from minimising transport time [33,35] to ensuring fairness in the distribution of blood products [25]. Most of the publications include a case study [25,34,35,39,42,51,52,76,81,86–91,93] with most of these based on the scenario of an earthquake in either Istanbul, Turkey or various cities in Iran, whilst one focusses on blood collection during the COVID-19 outbreak in Iran [87]. There is a lack of research within this area from elsewhere in the world, particularly studying disasters with anthropogenic causes.

2.2.1.4 Donor Demographics

Donor demographics and behaviour have a significant impact on the success of the collections echelon, and while many articles indirectly consider these aspects, only seven directly incorporate such aspects into their models. Alfonso et al. consider donor behaviour such as generosity and availability to inform their location-allocation model [5] and simulation model [6]. Custer et al. [56] study donor demographics regarding likelihood of donor deferrals, while Testik et al. [96] study donor arrival patterns. The location of donors is considered in various location planning models [5, 35, 79] to inform where to locate clinics to ensure the required amount of blood will be collected. Lee & Cheng [56] classify disparities in donor behaviour to identify possible causes of decreasing donations and predict donors' intentions, while Nagurney & Dutta [65] consider donor satisfaction and the effect of this on competing blood organisations. Lastly, both Haeri et al. and Hosseini-Motlagh et al. [43, 46] consider the link between donor motivation and a sufficient supply of blood.

2.2.1.5 Location/Clinic Planning

Location and clinic planning is the most popular category within the selected literature. This is unsurprising since the success of a clinic depends heavily on the location, which is ideally easily accessible and within a given radius of a large amount of regular donors in order to meet the collection targets. The majority of articles in this category detail a location-allocation problem [5, 8, 9, 23, 25, 33–35, 39, 41, 42, 44, 46–48, 50–52, 70, 76–79, 81, 84–86, 88–91, 95, 105] i.e. deciding the location of mobile and/or fixed donation clinics, and often also attempt to minimise the costs involved with moving clinics and transportation of blood products. The next most popular area of research within this category is the planning of clinics, such as the explicit scheduling of clinics at locations that are already assigned [5, 85]. Several articles in this category either focus on relocating or establishing a new facility such as a blood centre or stock holding unit [16, 23, 50, 76, 84, 93], whilst the comparison of effectiveness of mobile and fixed clinics is also considered [6]. Centralisation of a regional blood service is discussed in the literature, with various levels of centralisation considered and analysed by Osorio et al. [70]. Finally, a variety of clinic configurations are considered by Doneda et al. [30] i.e. clinic layout and number of donation chairs.

2.2.1.6 Staff Utilisation

Staff Utilisation is another category lacking in research, with only seven articles exploring this area. All of these propose slightly different approaches, but nearly all deal with determining the general staffing requirements for donation clinics in order for the blood services to reach their targets for donor satisfaction and volume of collected blood. Alfonso et al. [6], Doneda e al. [30] and Blake & Shimla [17] all consider various configurations of clinic staff and the impact of these on donor waiting times and service level, with the former two studies using simulation and the latter using queueing theory. The aim for all three models is to keep costs and queues to a minimum, and along with Testik et al. [96], intend to better inform policy. However, Testik et al. approach the problem from the perspective of donor arrivals and the effect of this on workforce utilisation. The goal of this is to identify patterns in donor behaviour (through data mining methods) and determine an adaptive workforce with varying numbers of staff throughout the day to better cope with changes in donor arrivals. Li et al. [58] consider staff schedules for blood donation clinics across Beijing, focussing on meeting donor demand by scheduling staff to transfer to another clinic in the same working day.

Van Brummelen [98] presents an ILP model to optimally assign varying length shifts to staff to best cope with donor arrival patterns. Van Brummelen also considers intra-day scheduling of staff across the stations within the clinic to minimise donor waiting times and staff hours worked. This model is based on fixed clinic sites, whereas Alfonso et al. [5] consider staff scheduling for mobile sites also. Alfonso et al. [5] present a two-stage model in which the latter produces a staff schedule for each clinic. However, the model does not include intra-day scheduling. Furthermore, none of the articles consider intra-day scheduling regarding employee breaks or closure of clinic for lunch and how this may effect donor flow. This lack of research in explicit scheduling of staff at donation clinics is surprising, since a lack of efficient staffing can have a massive impact on donor satisfaction, volume of collected blood, and monetary costs.

2.2.1.7 Vehicle Routing

In all publications under this category, the principle aim is to minimise the distance travelled by vehicles associated with the collections echelon, though the motivation behind this aim varies. Several articles discuss the importance of vehicle routing in order to maximise platelet production [63, 73, 77] due to its perishable nature. Şahinyazan et al. [85] and Gunpinar & Centeno [41] both utilise vehicle routing in order to maximise the number of blood donations collected, while Lodree et al. [59] and Razavi et al. [81] determine the optimal routes during the response phase following a large-scale disaster.

2.2.2 Methods

Since this thesis utilises mathematical modelling methods, a detailed overview about the methods and solution approaches of relevant literature is given in Table 2.5. The approaches are categorised into Data Mining and Machine Learning, General Statistical Analysis, Game Theory, Goal Programming, Heuristics, Integer Programming (includes mixed integer programming), Queueing Theory, Qualitative, Simulation, and Stochastic Modelling. Note that any previous literature reviews are assigned to the Qualitative category, and that the category Statistical Analysis includes articles that incorporate probability distributions or forecasts of certain aspects such as blood product demand or donor behaviour.

Table 2.5: Methods

Methods	**Articles**
Data Mining and Machine Learning	[52, 56, 63, 73, 96]
Statistical Analysis	[5, 27, 29, 60, 61]
Game Theory	[65]
Goal Programming	[23, 42, 43, 52, 81]
Heuristics	[48, 59, 60, 63, 70, 72, 73, 77, 78, 81, 85, 93, 95, 98]
Integer Programming	[5, 8, 9, 13, 22, 24, 25, 34, 35, 39, 41, 44, 46, 47, 50, 51, 58, 59, 63, 69, 71, 72, 76, 77, 79, 84, 85, 88, 90, 91, 93, 95, 98, 103–105]
Queueing Theory	[17, 96, 98]
Qualitative	[11, 12, 14, 61, 68, 75, 90]
Simulation	[6, 16, 27, 30, 60–62, 71, 86, 98, 102]
Stochastic Modelling	[9, 10, 25, 33, 35, 39, 41–44, 46, 47, 58, 69, 71, 77, 79, 86–89, 103–105]

The table reveals that (mixed) integer programming is a commonly used modelling and solution method in the field of blood collections. All of the articles that use this method, utilise it to solve an allocation problem: either of location, appointments, staff shifts, or routes of vehicles, with the objective function mostly being to minimise costs and to maximise donor satisfaction or blood collected. This conveys why integer programming is the most popular approach, since the effective allocation of clinic locations, donor appointments and vehicle routes are fundamental to the success of a blood supply chain.

Stochastic modelling is also a popular method in optimisation of collection of blood, and mostly used alongside an integer programme [9, 25, 35, 41, 44, 46, 47, 58, 69, 71, 77, 79, 88, 103–105]. This is due to the stochastic nature of the blood supply chain, particularly regarding both the demand of blood products and supply from donors. A significant amount of these publications present a robust optimisation approach [9, 25, 35, 41–43, 46, 47, 58, 79, 86, 89, 104, 105] which mostly consider the uncertainty in parameters such as donor arrival, costs and demand. For instance, Zahiri et al. [105] propose a strategic robust possibilistic programming model to ensure the results are still relevant and applicable over the long planning horizon, minimising the effect of changes in parameters over time. Many of these publications focus on disaster relief [25, 33, 35, 39, 42, 46, 47, 86–89] and utilise stochastic modelling to create a more robust blood supply chain, to ensure effectiveness even in the midst of an emergency.

Heuristic techniques involve algorithms that seek approximate solutions quickly, and these are a frequently used mathematical method within the selected literature. These are also often used alongside integer programming methods [59, 63, 72, 77, 85, 93, 95], and in this case are mostly integer-programming-based algorithms i.e. the relaxed linear programme solution is utilised as a construction heuristic. Heuristics are also used with statistical methods to aid a primarily simulation approach [60], as Lowalekar et al. use a gradient search-based heuristic to identify the optimum policy parameters for their model. Van Brummelen [98] uses a heuristic algorithm to allocate appropriate appointment slots throughout the day. Only one of these articles uses heuristics as their primary method [48] - Hsieh et al. propose a solution to a location-allocation problem regarding donation clinics, and here they use a sorting genetic algorithm to search for the Pareto set to solve the multi-objective problem. A common theme in the utilisation of heuristic algorithms is vehicle routing problems [59, 73, 77, 81, 85] due to the complexity associated with such problems, as they are NP-complete and cannot be solved exactly. The efficiency of heuristic methods is beneficial to many other problems within the blood supply chain, as such problems are often very complex with a large amount of parameters and variables.

Simulation is widely used to analyse and optimise the blood collections process. However, some of the articles in this category only use simulation as support or stochastic evaluation of the mathematical model, and it is therefore not the primary method in use. For example, simulation is used as a way of evaluating the implementation of a proposed model [60, 61, 86, 98], evaluating a current systems performance [6], or to generate scenarios for a mathematical programming formulation [5]. The only articles in which simulation is used as the primary method to optimise blood collections are [16, 27, 30, 62, 102]. Blake et al. [16] use simulation to determine the impact of the addition of a stock holding unit in a given region in

Canada, Lowalekar & Ravichandran [62] use simulation to compare two potential new collection policies against each other and indeed against the current policy. Custer et al. [27] evaluate the cost of blood per unit using simulation, while Doneda et al. [30] use simulation to observe the effect on donor time in the system and cost of a clinic under several different clinic configurations. Finally, Xu et al. [102] utilise simulation to compare different inter-donation intervals among male donors.

Data mining and machine learning techniques are mainly used to support other mathematical methods. For instance, Mobasher et al. [63] and Özener & Ekici [73] utilise clustering algorithms to assist with vehicle routing problems. Meanwhile, Testik et al. [96] and Lee & Cheng [56] use data mining and clustering methods to evaluate donor behaviour such as likelihood to donate and arrival patterns at clinics. Lastly, Khalilpourazari et al. [52] utilise neural learning methods to gain knowledge from past experiences to adapt to new challenges such as natural disasters. These methods provide an innovative approach to modelling of the blood collections process as they offer the opportunity for donation clinics to be planned in alignment with the respective donor-base of a given region, regarding the planning of location, capacity and staff.

2.2.3 Model Objectives

In what follows, we will break down articles which provide a mathematical model of the blood collection problem into various objectives; such objectives are clearly identifiable in the case of some methods, such as integer programming. For other methods, we identify objectives from the authors detail on the aims of their research, and outcomes of their model. A total of 54 articles are found to specify an objective to be maximised or minimised, with the most popular of these detailed in Table 2.6 whilst Table 2.7 displays the articles which share a common objective. The five objectives are defined as the following:

- **Maximise blood collection:** Models that aim to collect as much blood as possible.
- **Minimise blood shortage:** Any models that seek to minimise the shortage of blood products (less than demand).
- **Minimise cost:** Models that aim to lessen any costs, from clinic operation to transportation costs
- **Minimise distance:** Any models that aim to decrease distance in some way, ranging from distance travelled to distance between facilities.
- **Minimise time:** This includes the minimisation of time used - either regarding a specific aspect (such as transportation), or that of the whole blood supply chain.

Both Table 2.6 and Table 2.7 clearly show that the most popular objective is minimisation of cost. This is unsurprising since most blood services are non-profit organisations, and donors are usually voluntary and non-remunerated. While many articles aim to minimise the cost of the collection echelon in general, or even the whole blood supply chain, some focus on more specific costs. An example of this is the cost of clinic operation; Salehi et al. [86] seek to minimise the costs of establishing permanent blood centres, while Blake & Shimla [17] aim to minimise costs associated with staffing.

The minimisation of time is the second most popular objective, which is typically due to the perishable nature of blood. This objective is often considered alongside minimisation of costs as the two are closely linked, especially regarding transportation and staffing. Some articles present a model which aims to minimise the length of the blood supply chain, across all echelons i.e. reduce the time blood products spend in the system, from donation to distribution at a hospital. For example, Attari et al. and Arvan et al. [8, 9] present this goal in their respective models.

Table 2.6: Popular Objectives

Objective	**Articles**
Maximise Blood Collection	[35, 58, 63, 77, 85]
Minimise Blood Shortage	[25, 34, 39, 42, 52, 61, 71, 78, 91]
Minimise Cost	[8–10, 17, 22–25, 30, 33–35, 39, 42–44, 46, 47, 50–52, 58, 65, 69–72, 76–79, 81, 84–91, 93, 95, 98, 104, 105]
Minimise Distance	[23, 24, 41, 86]
Minimise Time	[8, 9, 30, 33, 35, 39, 44, 51, 52, 73, 84, 87, 95]

Table 2.7: Common Objectives

Objectives	**Articles**																							
	[8]	**[9]**	**[23]**	**[24]**	**[25]**	**[30]**	**[85]**	**[33]**	**[34]**	**[35]**	**[84]**	**[39]**	**[42]**	**[44]**	**[51]**	**[52]**	**[58]**	**[71]**	**[77]**	**[78]**	**[86]**	**[87]**	**[91]**	**[95]**
Maximise Blood Collection							✓			✓							✓		✓					
Minimise Blood Shortage					✓				✓			✓	✓					✓		✓			✓	
Minimise Cost	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Minimise Distance			✓	✓																	✓			
Minimise Time	✓	✓				✓		✓		✓	✓	✓		✓	✓	✓						✓		✓

However, other models focus on a specific attribute such as transportation time [33,35,39,51,52,73,95] - distance is closely related to this, which may be the reasoning behind fewer articles detailing the minimisation of distance as a main objective. Although, minimising distance may also concern factors other than transportation, such as fairness in distances between blood centres and hospitals within a given region [23]. Additionally, Şahin et al. [84] aim to minimise the total demand-weighted distances both from donation clinics to blood processing centres, and from blood processing centres to hospitals.

Five articles detail the maximisation of blood collected as an objective, though from varying perspectives. Rabbani et al. [77] and Mobasher et al. [63] wish to maximise the amount of donations that can be processed for platelet production, while Şahinyazan et al. [85] aim to maximise the amount of blood collected while optimising vehicle routing between mobile clinics. Fazil-Khalaf et al. [35] seek to collect as much blood as possible in disaster situations, while Li et al. [58] aim to collect as much blood as possible whilst reducing staff transfer costs. Though each blood service requires a significant amount of blood to support demand, perhaps this objective is less popular due to overcollection of blood leading to wastage of blood products.

A total of nine articles explicitly aim to minimise the shortage of blood products i.e. the amount of viable donations less than the demand. Matching supply to demand is the fundamental goal of blood services as not meeting the blood product demand of patients has fatal consequences. Thus, the number of articles considering this as an objective is surprisingly low, and indicates a possible priority for future research in the area of optimisation of the blood collection echelon. However, some articles utilise a constraint ensuring that the blood demand is met as a constraint and therefore would not appear as an objective function.

Aside from the objectives listed in Table 2.7, various objectives were discussed individual to specific articles. For example, since the functional area of [13] is appointment scheduling, in this article, Baş et al. focus their objective function on balancing the production of each blood type among days to minimise wastage of blood. Alfonso et al. [5] aim to minimise the system overtime, whilst maximising the donor service level. Mobasher et al. [63] focus on minimising the total working time along with the blood supplied by other regions (and thus incurring a cost). Lowalekar & Ravi [61] concentrate on inventory-related objectives and aim to minimise both shortages and outdates of blood products, whilst Testik et al. [96] focus on clinic operation related objectives such as maximisation of staff utilisation and minimising donor wait.

A wide variety of objectives are covered in the selected literature, and whilst minimising costs is of high importance to a blood service, donor satisfaction is crucial and often overlooked. As discussed above, the minimisation of donor waiting times and service level has been considered, though minimally. The matching of supply and demand is vital to the success and effectiveness of a blood supply chain, and there is little optimisation of this considered in the literature, as it is usually indirectly - this implies that further research which places matching supply and demand (not simply meeting demand) as a main objective would be of great benefit to blood supply chains worldwide and reduce both 'overcollection' and wastage of blood.

2.2.4 Planning Decision Levels

The selected articles can be categorised by the planning decision level that they discuss; namely *strategic, tactical or operational (offline or online).*

Table 2.8: Articles by Planning Decision Levels

Planning Decision Level	**Articles**
Strategic	[5, 8, 9, 16, 22–25, 27, 29, 33–35, 39, 42, 43, 46, 47, 50–52, 56, 60, 61, 70, 71, 78, 79, 84, 86, 88–91, 95, 104, 105]
Tactical	[9, 10, 13, 30, 35, 39, 44, 46, 48, 52, 58, 59, 63, 65, 69, 72, 73, 76, 77, 81, 85, 87–89, 93, 98, 102, 103, 105]
Operational	
Offline	[5, 6, 13, 17, 41, 58, 63, 71, 85, 96, 98]
Online	[13, 98]

As described by Hulshof et al. [49], strategic planning ‘addresses structural decision making’ and involves the decisions which help to develop and improve an organisation, more specifically in our case, a blood service. This is therefore typically over a long planning horizon. As shown in Table 2.8, a significant number of the articles deal with strategic planning decisions. This is due to the vast number of articles that deal with locational planning and collection policy, as changes to these are typically implemented incrementally, and over a long period of time.

Tactical planning ‘translates strategic planning decisions to guidelines that facilitate operational planning decisions’, and often involves the coordination of operations within an organisation, as described by Hulshof et al. [49]. These decisions essentially focus on the ‘what, where, how, when and who’ of a given process. A total of 29 of the selected articles fit into this category, and these are mainly either appointment scheduling,vehicle routing problems, or insights about how to adapt during an emergency - these types of problems are often solved by providing a framework for the blood service to implement, a decision support tool, or specific managerial insights.

Operational planning is typically on a short-term basis, involving the execution of a blood service’s processes; this planning decision level is further categorised into offline and online planning. Operational offline decisions are those that are made in advance of a process being carried out, such as assigning a resource to a donor, while operational online decisions are ‘control mechanisms that deal with . . . reacting to unplanned events’ [49] during the process. All of the articles in the operational planning category are also in the operational offline planning subcategory, with only two also being in the operational online planning subcategory. This illustrates the need for processes to be well-prepared and organised in advance of starting a donation clinic, but also online decisions may need to be made regarding sickness of staff, appointment cancellations and prioritisation of donors.

Due to the complex nature of the blood supply chain, all planning decision levels are important and necessary. However, strategic planning allows for perhaps the most significant improvements towards a more effective and efficient supply chain, as long-term goals are able to be realised, such as the alignment of demand and supply.

2.2.5 Case Studies and Implementation

A total of 51 of the articles (75%) include a case study, with real-life data from a chosen blood service, as seen in Table 2.9. Several of these articles are of research motivated by a specific blood service, with the aim of any findings being implemented if proven effective, or to inform future policy.

Table 2.9: Articles by Case Study Inclusion

Case Study Included	**Articles**
Yes	[5, 6, 9, 10, 13, 16, 17, 22, 25, 27, 29, 30, 34, 35, 39, 41–44, 46–48, 50–52, 60–63, 65, 69, 70, 72, 76, 79, 81, 84–91, 93, 96, 98, 102–105]
No	[8, 11, 12, 14, 23, 24, 33, 56, 58, 59, 68, 71, 73, 75, 77, 78, 95]

Only 17 of the articles did not present a case study within their research - note that the five previous literature reviews are included in this category, and aside from these, the remaining articles propose findings which are either tested on hypothetical data/situations or simply presented as theoretical arguments.

The majority of selected publications presented a solution which could possibly be implemented in real-life scenarios, yet only five of these articles specifically stated that their proposal had either been implemented or are likely to be implemented [10,22,50,71,103]. These findings are similar to that of Brailsford et al. [21] who state that 'levels of implementation for models in healthcare OR are very small indeed'. Harper and Pitt [45] discuss common reasons for this lack of implementation in healthcare settings. However, this figure may not be a true representation of the implementation of the current research, due to the possibility that the timeline of many projects may have ended before implementation could be carried out, and thus the corresponding article would have no mention of such.

2.3 Literature Review Findings

This chapter has presented a categorisation framework to distinguish methods, functional areas and various other aspects of existing research, and provides a clear classification of the literature surrounding blood collections. This enabled us to identify any areas that require further research.

Perhaps the most notable area which calls for more in-depth research to be conducted is that of resource planning at blood donation clinics, and more specifically workforce planning. Presently, very few publications explore this (within our literature search) with most simply analysing the levels of required staff instead of explicitly scheduling staff optimally. Van Brummelen [98] does consider intra-day scheduling of staff and varying shift lengths but only for fixed clinic sites. Only one model explicitly generates a staff schedule for mobile donation clinics [5], but this does not consider intra-day scheduling. Not only does the assignment of staff to clinics require further research, but also the scheduling of staff throughout a day, especially considering staff breaks and how to mitigate the effects of this on the donor waiting times. Intra-day scheduling is of great importance as it can help to improve donor service level, reduce waiting times and increase productivity, through utilising the workforce in alignment with varying donor behaviour throughout the day.

Appointment scheduling is also lacking in research within optimisation of the collection of blood, despite appointments enabling clinics to have some control over queues and donor arrivals. In the existing literature, the research tends to focus on one aspect of appointment scheduling such as apheresis donations, blood types, or aligning appointments with transportation of collected units. However, appointment scheduling provides the opportunity to not only control inventory in regards to volume and blood type, but also to manage the donor flow through a clinic via analysis of frequency of appointment slots. There is much research to be undertaken that marries all of these important aspects together.

Both the scheduling of staff and appointments have great impact on the efficiency of a clinic and donor satisfaction. Due to donors (in most cases) being volunteers and non-remunerated, it is vital to maximise the experience of donors as far as possible i.e. minimisation of queues, efficient service, convenient location and appointment time, etc. Future research within this field should place donor satisfaction at the forefront of its aims and objectives, as the success of blood supply chains ultimately rely on the many generous donors worldwide.

Matching supply and demand is the goal of any supply chain, but achieving this goal is of utmost importance to the blood supply chain, with failure to do so resulting in potentially critical consequences. Despite this, matching the supply and demand of blood is often neglected or only indirectly considered. Whilst many publications account for demand being satisfied, there is a significant lack of research regarding 'overcollection' of blood as this leads to avoidable wastage of invaluable products.

In conclusion, as human blood is an invaluable and scarce resource which is essential to modern healthcare, the blood supply chain is vitally important globally. Since the success of the blood supply chain is ultimately dependent on voluntary donors in most parts of the world, further research into the collection of donated blood is imperative to reduce wastage and shortages of blood, and increase the effectiveness and efficiency of blood services.

2.4 Summary

This literature review chapter has focussed primarily on articles which discuss the analysis and optimisation of the collections echelon of the blood supply chain, as there is a distinct lack of existing reviews providing an extensive and recent evaluation of this particular field of research. Gaps in the research area have been identified; namely the optimisation of staff scheduling including intra-day scheduling, appointment scheduling for donors, and the direct matching of supply to demand for blood products.

In the following chapter, Chapter 3, the Welsh Blood Service collection model is described and the limitations of this current practice are identified. Further, the Blood Donation Clinic Scheduling Model is presented in a mathematical formulation in the format of a linear programme, to optimally schedule clinics across a given planning horizon. Decision variables are described, along with all other parameters, before three alternative objective functions and constraints are presented.

Chapter 3

The Blood Donation Clinic Scheduling Problem

This chapter describes the current practice for scheduling blood donation clinics at the Welsh Blood Service (WBS), the limitations of this practice, and proposes a mathematical model to improve the efficiency of the clinic scheduling process. This chapter is structured as follows: the current practice of blood donation clinic scheduling at the WBS is described in Section 3.1, with the mathematical formulation of the Blood Donation Clinic Scheduling Problem (BDCSP) introduced in the subsequent Section 3.2 with all relevant parameters. The Blood Donation Clinic Scheduling Model (BDCSM) formulation is presented in Section 3.3 with a description of all decision variables, objective functions and constraints.

3.1 Current Practice

It has been identified that the way in which the WBS collects their blood donations, though consistently meeting demand for blood products, would benefit from an increase in efficiency. At present, the WBS often exceeds their target of number of blood units collected, resulting in the demand being satisfied but also increasing the likelihood of wastage of product due to the perishable nature of blood products. This then renders the donors of any wasted product ineligible for donation for a minimum of 12 weeks for male donors and 16 weeks for female donors[1]. This also implies that less clinic hours are required to meet the demand (within some tolerance to account for variation between predicted and actual demand). In addition, the Welsh Government require that the WBS waste no more than 2% of all collected units each calendar year, with potential fines presented if the wastage exceeds this. Therefore, matching supply to demand is of great importance to the continued success of the Welsh Blood Service and indeed the blood supply chain in Wales.

The Welsh Blood Service collects nearly all of their whole blood donations from 'mobile' clinics, i.e. temporary clinics, with only one permanent clinic based at their headquarters in Talbot Green. Every temporary clinic is either located in a venue (often hired at cost) or in a parked trailer (owned by the WBS) at a specific location. The former of these clinic types are referred to as 'community' clinics and the latter as 'trailer' clinics. As of 2016, the WBS collects blood donations from all counties in Wales and is the only organisation operating blood donation clinics in Wales.

3.1.1 Regions

The WBS has four 'bases' i.e. locations where clinic-based workers start their shift to travel to a donation clinic, and also where equipment and vehicles are stored. Every clinic is assigned to a specific base such that only resources from the assigned base serve this clinic. This allows Wales to be split into four regions operationally with each base as the hub of each region, as displayed in Figure 1.2.

The South East region has significantly more clinics than any of the other regions due to the population density being much larger in this region with almost half of the Welsh population residing in this area [40], and therefore there are more donors to reach. For this reason, the WBS has the resources to operate up to four

[1] Where male and female donors are referring to those assigned the respective gender at birth.

clinics a day in the South East region. Each other region (South West, North West and North East) has only one team of clinic workers and therefore the capacity to operate at most one clinic per day in each region.

3.1.2 Clinic Venue Types

There are two clinic venue types (excluding the one permanent clinic at the WBS headquarters) namely 'community' clinics and 'trailer' clinics. The former usually take place in a venue such as a local community centre whereas the latter usually take place in a parked trailer (often outside a supermarket). Community clinics typically have a larger capacity than trailer clinics, as the latter are often limited to a maximum of 6 donation chairs. The number of trailer clinics that can operate on a given day is limited as there are a fixed number of them available to the WBS.

In addition, some clinics are organised with a given company to operate at a particular workplace, with appointments to be reserved for employees only, and therefore closed to the general public. These clinics are called 'company' clinics, with all other clinics deemed to be 'public' clinics.

3.1.3 Clinic Tours

Due to the bases being located either far to the north or south of Wales, Mid-Wales clinics such as Aberystwyth involve significant travel time. In order to serve these more rural areas of Wales, as obligated by the Welsh Government, these clinics are operated via clinic tours. A clinic tour consists of a team being deployed to work multiple consecutive days at clinics in Mid-Wales, and staying overnight in the local area to minimise travel time. The tours are pre-determined with specific clinics grouped together in a set order to further minimise travel time. For example, the South East base serves the tour illustrated by Table 3.1.

In this specific tour, workers leave the South East base on the morning of day one, travel to the first location of Rhayader and set up the clinic in the hired venue. At

Table 3.1: Example of a Clinic Tour

Day	1	2	3	4
Clinic Location	Rhayader	Builth Wells	Llandrindod Wells	Llandrindod Wells

the end of the clinic, all equipment is loaded onto the WBS lorry and the workers travel to nearby booked accommodation. Similarly, on the second day, the workers travel from the booked accommodation of the first night to Builth Wells to set up the clinic there and so on, until at the end of the fourth day, the workers travel back to the South East base. Donated blood products are collected each day from the scheduled tour clinic and transported back to the processing centre separately.

3.1.4 Clinic Frequency

Due to the eligibility requirements regarding frequency of donation i.e. males must wait at least 12 weeks and females must wait at least 16 weeks between each donation, clinics are limited to how often they can operate in the same location. In most cases, a given clinic will not be scheduled to operate more frequently than every 16 weeks. However, if the donor panel of a clinic is large enough to support a given clinic being scheduled more frequently, then it may be scheduled every six weeks, or three times within every 16 week period. This only occurs in areas of higher population density.

The Welsh Government require the WBS to give every eligible person in Wales the option to donate blood i.e. donors should not have to travel more than a reasonable distance to a donation clinic. This means that more rural clinics such as those in Mid-Wales, which generally have a very small donor pool and therefore return only a small amount of whole blood donations, must be scheduled at least once every calendar year.

3.1.5 Clinic Availability

Each clinic has an associated weekday availability pattern and a seasonal availability pattern due to the variety of venues and their requirements and/or schedules. A weekday availability pattern details the specific weekdays that a given clinic is available e.g. 'weekdays only' or 'Mondays only'. Similarly, a seasonal pattern details the availability of a given clinic over the calendar year, e.g. 'March - October only' or 'school holidays only'.

3.1.6 Clinic Duration Pattern

Each donation clinic has an associated duration pattern i.e. the pattern of days that a given clinic is to run each time it is scheduled. The vast majority of the WBS clinics have a duration of one day, and therefore start and end on the same day. However, a minority of clinics run for a number of days, usually consecutive, with some clinics having more complicated duration patterns such as three consecutive days one week followed by two consecutive days the next week.

3.1.7 Blood Supply

Each clinic has a capacity of donors each day that it operates based on the physical capacity of the venue, the number of hours that the clinic is in operation, and the number of workers available. Therefore, each clinic has a maximum supply of whole blood that can be collected each day that it runs. Due to eligibility screening in the clinic, failed venepuncture, and 'Do Not Attends' (DNAs) i.e. donors failing to attend their appointment, the maximum supply of a clinic is unlikely to be met.

The WBS typically calculates their estimated supply per clinic based on an average bleed figure over the last three times the given clinic occurred, where bleed figures relate to the number of donors that provided a blood donation (viable or not). The WBS utilise these supply estimate figures during clinic scheduling to ensure the demand is met.

3.1.8 Demand

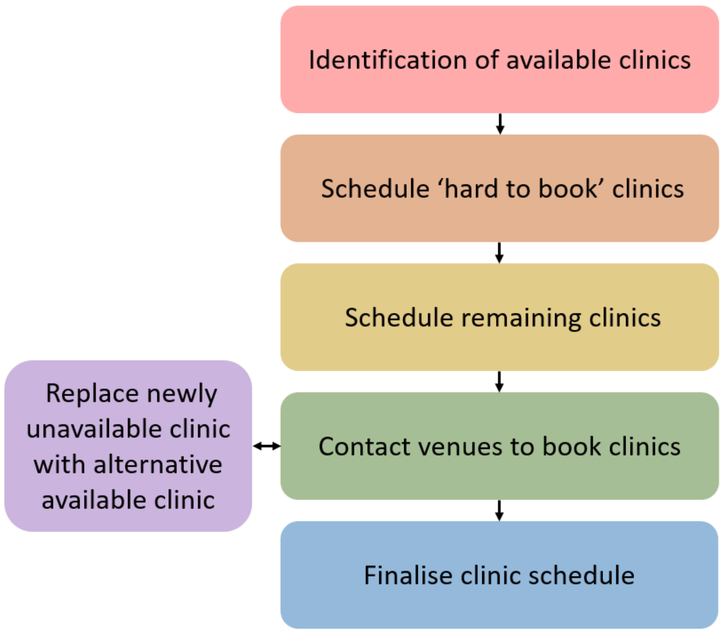
The Welsh Blood Service delivers blood products to all hospitals across Wales, both NHS and private sector. All collected blood products are processed at the laboratory at the WBS headquarters in the South East region and delivered to hospitals as required.

The WBS is undertaking research to identify the 'true demand', as supplied hospitals work independently from the WBS and are in control of determining their own blood product orders. This means that at present the actual demand is not known as the hospitals do not report back their usage. However, throughout this research, we have assumed the demand to be that of meeting the hospitals' orders.

3.1.9 Clinic Scheduling Process

Currently, the WBS schedule their clinics using a manual and tedious process, working on four-week planning periods at a time. The process is described in Figure 3.1.

Figure 3.1: Welsh Blood Service Clinic Scheduling Process



The first stage of the process shown in Figure 3.1 involves the clinic planner identifying clinics which are available to be scheduled during the given planning horizon by assessing the date of the previous time each clinic was scheduled against the maximum frequency of each clinic. In most cases, the maximum frequency is 16 weeks (due to donor eligibilty) but this may be as often as every four weeks if a clinic is in a particularly densely populated area and therefore has a large donor panel to support this frequency. Following this, the clinic planner prioritises clinics that are listed as 'hard to book' (this is often due to the venue having limited availability) and clinics in this category need to be booked much further in advance than others.

Once any 'hard to book' clinics are scheduled within the given planning horizon, the other clinics can then begin to be scheduled. These clinics are scheduled with the requirement to meet the estimated weekly demand (using the estimated collection figure per clinic and also considering the availability pattern of each clinic) but also ensuring that staff contractual hours are met. This often results in more clinics being operated than is necessary to meet demand. Once the first draft of the schedule is complete, the associated venues are contacted to make a booking on the required day(s). If an issue arises during this process i.e. a venue is not available for the allocated day(s), the schedule is iteratively changed; an alternative available clinic is selected to replace the newly unavailable clinic. This last process is repeated until all clinic bookings are confirmed and secured.

3.1.10 Limitations of Current Practice

The current clinic scheduling practice and blood collection strategy at the WBS causes various inefficiencies to arise. Firstly, the current aim of the WBS is for each four-week clinic schedule to not only to meet the estimated blood product demand, but also to satisfy the contractual hours of the clinic-based workforce. This likely results in unnecessary monetary costs, including transportation and venue hire costs, as a result of potentially operating more clinics than is required to meet demand. This also implies that there are more contractual hours for the clinic-based workforce than required. This is the motivation behind objective function one (3.3) described in Section 3.3.2 i.e. the minimisation of the number of clinic days scheduled over a given planning horizon.

In addition to this, if the overcollection rate is consistently higher than is desired (with some tolerance needed to counteract variation in demand) the production of these blood products may end in waste, whilst the donors themselves remain ineligible to donate until 12 (male) or 16 (female) weeks have passed. This is the motivation behind objective function two (3.4) - the minimisation of overcollection i.e. the minimisation of estimated collected blood donations in excess of the demand.

Each four-week clinic schedule is created manually which poses a complex task for clinic planning team, with many restrictions to consider. Most of the constraints associated with a clinic are logged digitally, though these are often across multiple files and are not in one central location, with some constraints dependent on the relay of local knowledge by the clinic planner(s). The intractability of the number of constraints to manage whilst creating a clinic schedule results in each initial schedule requiring approximately five working days to be completed by the clinic planning team. This schedule then may undergo various amendments in an iterative process to account for any venues that are newly-unavailable or other changes in circumstances until a schedule is finalised.

We aim to address these limitations in the following model by providing the WBS with a more streamlined and efficient clinic scheduling process that enables all restrictions and constraints to be considered whilst better matching supply to demand and reducing unnecessary costs by operating fewer clinics where possible.

3.2 The Blood Donation Clinic Scheduling Problem Formulation

Stage one of our mathematical model schedules blood donation clinics in alignment with the estimated supply at various locations and demand for blood. We formulate the problem as follows.

3.2.1 Planning horizon:

Let $\mathcal{T} = \{1, \ldots, T\}$ be the set of discrete times within the planning horizon of length T, such that T is a multiple of seven. This set of discrete times are equally incremented as days, and thus we will relate to each time increment as one day, $t \in \mathcal{T}$. We interchangeably refer to $\mathcal{T}$ and the planning horizon as one and the same. Each planning horizon is to begin on a Monday (aligned with how the WBS operates) with each full calendar week during the planning horizon within the set $\mathcal{W} = \{1, \ldots, T/7\}$ of weeks.

3.2.2 Clinics and Clinic-Dependent Parameters:

Set $\mathcal{I}$ depicts a set of established clinics, with each unique clinic denoted as i. We introduce tuples of clinics, (i_1, i_2), and a set of such tuples, $\mathcal{N}_b \subset \mathcal{I} \times \mathcal{I}$. Each tuple $(i_1, i_2) \in \mathcal{N}_b$ consists of clinics that cannot run within $b \in \mathbb{N}$ days of each other. Therefore, let b_{i_1,i_2} denote the number of days that must lapse between occurrences of clinics i_1 and i_2. Each clinic $i \in \mathcal{I}$ has the following attributes:

- **Clinic Duration Pattern:** Let $c_{i,t,t'} = 1$ if a clinic $i \in \mathcal{I}$ is scheduled to begin on day $t \in \mathcal{T}$ and therefore is also scheduled to run on day $t' \in \mathcal{T}$ due to the duration pattern of clinic i. Where clinic i has a duration of one day each time it is scheduled, clinic i can only run on the 'starting day' and thus $c_{i,t,t'} = 1$ if, and only if $t = t'$. However if clinic i occurs for a given number of consecutive days each time it is scheduled, for example three days, then $c_{i,t,t'} = 1$ where $t' \in \{t, t+1, t+2\}$. An example of a clinic duration pattern is illustrated in Table 3.2 below.

- **Frequency:** Let f_i be the minimum number of days that must lapse between each time clinic $i \in \mathcal{I}$ is scheduled. Some clinics are required to be scheduled much more frequently and as such must be scheduled within each planning horizon; Let $\mathcal{I}^{\text{forced}} \subset \mathcal{I}$ denote the set of clinics that are obligated to be scheduled within each planning horizon.

- **Availability:** Let $a_{i,t}^{\text{weekday}} = 1$ if clinic $i \in \mathcal{I}$ is available to be scheduled on day $t \in \mathcal{T}$ regarding weekday availability, and 0 otherwise. Let $a_{i,t}^{\text{seasonal}} = 1$ if clinic $i \in \mathcal{I}$ is available to be scheduled on day $t \in \mathcal{T}$ regarding seasonal availability, and 0 otherwise.

- **Estimated Supply:** Let $s_{i,t}$ represent the estimated supply of blood to be collected from clinic $i \in \mathcal{I}$ if it takes place on day $t \in \mathcal{T}$.

- **Region:** Let $r \in \mathcal{R}$ denote the set of regions. Let $\mathcal{I}^r$ denote the subset of clinics that belong to region r. Let $R_{t,r}^{max}$ denote the maximum number of clinics that can be scheduled to occur on day $t \in \mathcal{T}$ for region r. Similarly, let $R_{w,r}^{max}$ denote the maximum number of clinics that can be scheduled to occur over a calendar week $w \in \mathcal{W}$ in region r. Finally, let $R_{\mathcal{T},r}^{min}$ denote the minimum number of clinic days to be scheduled during the whole planning horizon $\mathcal{T}$ for region r.

- **Length of Day:** Let $l_{i,t}$ denote the expected duration of a working day (in hours) for clinic $i \in \mathcal{I}$ if it takes place on day $t \in \mathcal{T}$. Let $\mathcal{L}_{12}$ be the set of clinics i with length of day greater than or equal to 12 hours i.e. $l_{i,t} \geq 12$.

- **Venue Type:** Let i^{trailer} denote the clinics $i \in \mathcal{I}$ that take place in a trailer venue type. Therefore, $\mathcal{I}^{\text{trailer}}$ is the set of all trailer clinics.

- **Clinic Type:** Let $\mathcal{I}^{\text{tours}} \subset \mathcal{I}$ denote the set of tuples of clinics, $(i_1, ...i_K) \in \mathcal{I}^{\text{tours}}$ where each tuple forms a clinic tour. Let $k \in \mathbb{N}_{\leq K}$ represent each individual day of a tour, where $K \in \mathbb{N}$ denotes the final day number of a tour. Thus, K also denotes the length of a tour i.e. the number of consecutive days it requires each time it is scheduled.

- **Date Last Held:** Let o_i denote the date that clinic $i \in \mathcal{I}$ most recently occurred, prior to the current planning horizon $\mathcal{T}$.

- **Distance:** Let $\mathcal{Z}_j \subset \mathcal{I} \times \mathcal{I}$ denote the set of pairs of clinics (i_1, i_2) with a public donor panel type, that are within $j \in \mathbb{N}$ miles of each other.

Table 3.2: Example of a Clinic Duration Pattern

Monday	**Tuesday**	**Wednesday**	**Thursday**	**Friday**	**Saturday**	**Sunday**
1	1	1	0	0	0	0
0	1	1	1	0	0	0
0	0	1	1	1	0	0
0	0	0	1	1	1	0
0	0	0	0	1	1	1

**Example of a duration pattern of a clinic (for a one-week planning horizon) that operates for three consecutive days each time it is scheduled. If the clinic is selected to start on Monday, then it also runs Tuesday and Wednesday. However, it cannot start on a day later than Friday as it could not run for three consecutive days due to Sunday being the end of the planning horizon in this case.*

- **Number of Days Booked:** Let $g_i \in \mathbb{N}$ denote the number of days to be scheduled for clinic $i \in \mathcal{I}$, each time it is scheduled i.e. the number of clinic days for clinic i.

3.2.3 Demand

Let d_w be the total number of units of blood determined to be collected per week, $\forall w \in \mathcal{W}$, i.e. the weekly demand.

3.3 Model Formulation

With all parameters introduced, we now present the decision variables and their domains along with the constraints.

3.3.1 Decision variables

Let the binary decision variable, $x_{i,t}$ be introduced, such that;

$$x_{i,t} = \begin{cases} 1, & \text{if clinic } i \in \mathcal{I} \text{ starts on day } t \in \mathcal{T} \\ 0, & \text{otherwise} \end{cases} \tag{3.1}$$

Let $x_{i,t}$ denote the starting day of clinic $i \in \mathcal{I}$ i.e. the first day that a clinic runs

Table 3.3: Notation for Model Formulation

Sets	
$\mathcal{I}$	Set of clinics
$\mathcal{I}^r \subset \mathcal{I}$	Set of clinics in region $r \in \mathcal{R}$
$\mathcal{I}^{\text{trailer}} \subset \mathcal{I}$	Set of clinics of trailer venue type
$\mathcal{I}^{\text{tours}} \subset \mathcal{I}$	Set of clinic tours
$\mathcal{L}_{12}$	Set of clinics with $l_{i,t} \geq 12$
$\mathcal{N}_b \subset \mathcal{I} \times \mathcal{I}$	Set of pairs of clinics that must be at least $b \in \mathbb{N}$ days apart
$\mathcal{R}$	Set of regions
$\mathcal{T}$	Set of days
$\mathcal{W}$	Set of weeks
$\mathcal{Z}_j \subset \mathcal{I} \times \mathcal{I}$	Set of pairs of public clinics within $j \in \mathbb{N}$ miles of each other
Parameters	
$a_{i,t,}^{\text{weekday}}$	Weekday availability of clinic $i \in \mathcal{I}$ on day $t \in \mathcal{T}$
$a_{i,t}^{\text{seasonal}}$	Seasonal availability of clinic $i \in \mathcal{I}$ on day $t \in \mathcal{T}$
b_{i_1,i_2}	Number of days that must lapse between clinics $i_1 \in \mathcal{I}$ and $i_2 \in \mathcal{I}$
$c_{i,t,t'}$	Duration pattern of clinic $i \in \mathcal{I}$ on day $t' \in \mathcal{T}$ with starting day $t \in \mathcal{T}$
d_w	Blood collection demand for week $w \in \mathcal{W}$
f_i	Number of days that must lapse between consecutive occurrences of clinic $i \in \mathcal{I}$
g_i	Number of days to be scheduled for clinic $i \in \mathcal{I}$
$(i_1, ..., i_K) \in \mathcal{I}^{\text{tours}}$	Clinic tour consisting of $K \in \mathbb{N}$ clinics
$l_{i,t}$	Length of day for clinic $i \in \mathcal{I}$ on day $t \in \mathcal{T}$
o_i	Date that clinic $i \in \mathcal{I}$ last occurred
$R_{t,r}^{max}$	Maximum number of clinics that can be scheduled on day $t \in \mathcal{T}$ in region $r \in \mathcal{R}$
$R_{w,r}^{max}$	Maximum number of clinics that can be scheduled over week $w \in \mathcal{W}$ in region $r \in \mathcal{R}$
$R_{\mathcal{T},r}^{min}$	Minimum number of clinics to be scheduled over $\mathcal{T}$ in region $r \in \mathcal{R}$
$s_{i,t}$	Estimated supply at clinic $i \in \mathcal{I}$ on day $t \in \mathcal{T}$
$T \in \mathcal{T}$	Final day of the planning horizon $\mathcal{T}$
Decision Variables	
Δ_w^+	Units of blood collected in excess of demand in week $w \in \mathcal{W}$
$x_{i,t}$	1, if clinic $i \in \mathcal{I}$ starts on day $t \in \mathcal{T}$, 0 otherwise

each time it is scheduled, as clinics may have a duration that exceeds one day. The ‘stop’ day of a clinic is dependent on the clinic duration pattern, $c_{i,t,t'}$.

Let Δ_w^+ be the estimated total units of blood to be collected that exceeds the demand for each week $w \in \mathcal{W}$, i.e. the weekly amount overcollected such that;

$$\Delta_w^+ = \left(\sum_{i \in \mathcal{I}} \sum_{t \in \mathcal{T}_w} \sum_{t' \in \mathcal{T}_w} c_{i,t,t'} \cdot x_{i,t} \cdot s_{i,t} \right) - d_w \quad \forall\, w \in \mathcal{W} \tag{3.2}$$

3.3.2 Objective Function

Having introduced all sets, indices and decision variables we formulate the three alternative objectives as follows:

Objective Function One

$$\text{Minimise} \quad \sum_{i \in \mathcal{I}} \sum_{t \in \mathcal{T}} \sum_{t' \in \mathcal{T}} c_{i,t,t'} \cdot x_{i,t} \tag{3.3}$$

Objective Function Two

$$\text{Minimise} \quad \sum_{w \in \mathcal{W}} \Delta_w^+ \tag{3.4}$$

Objective Function Three

$$\text{Minimise} \quad \left(\sum_{i \in \mathcal{I}} \sum_{t \in \mathcal{T}} \sum_{t' \in \mathcal{T}} c_{i,t,t'} \cdot x_{i,t} \right) + \left(\sum_{w \in \mathcal{W}} \Delta_w^+ \right) \tag{3.5}$$

Objective function one (3.3) minimises the total number of scheduled clinic days over the planning horizon while objective function two (3.4) minimises the weekly overcollection for each week of the given planning horizon. Objective function three (3.5) combines both objective functions one and two and minimises both the number of clinics days scheduled and the weekly overcollection over the given planning horizon.

There are no weights currently assigned to the terms in objective function three (3.5), due to experimental results indicating that the values for both the total number of clinic days scheduled and the estimated overcollection (over the given planning horizon) were satisfactory. For nearly all instances utilising objective function three, the estimated overcollection was successfully minimised to zero, with the number of clinic days scheduled being the minimum permitted by the constraints included, or otherwise the minimum required to meet the demand. The only instances where this is not the case are the North East instances, due to the lack of options of clinics to schedule to achieve zero overcollection, and only one iteration (of a total of 10) for instance SW213, where one additional clinic is scheduled in place of one unit of overcollection.

It would be required that the WBS determine their preferences in terms of balancing the two variables (total clinic days and overcollection) if the model were to be implemented, to ensure that any solution meets their needs and priorities. One popular method to help determine how these weightings should be balanced is Pareto optimisation. We discuss the possible use of this method in Section 7.4.1.

3.3.3 Constraints

Demand Satisfaction Constraints

We establish the connection between demand, estimated supply and overcollection with the following constraint (3.6).

$$d_w - \left(\sum_{i \in \mathcal{I}} \sum_{t \in \mathcal{T}_w} \sum_{t' \in \mathcal{T}_w} c_{i,t,t'} \cdot x_{i,t} \cdot s_{i,t} \right) + \Delta_w^+ = 0 \qquad \forall w \in \mathcal{W} \tag{3.6}$$

The following constraints are to ensure that the estimated supply of each week in the planning horizon meets the weekly demand.

$$\sum_{i \in \mathcal{I}} \sum_{t \in \mathcal{T}_w} \sum_{t' \in \mathcal{T}_w} c_{i,t,t'} \cdot x_{i,t} \cdot s_{i,t'} \geq d_w \qquad \forall w \in \mathcal{W} \tag{3.7}$$

Frequency-based Availability Constraints

To ensure that the frequency-based availability of each clinic is considered, constraints (3.8) ensure that the minimum number of days have lapsed since the most recent time that the clinic was scheduled (prior to the planning horizon) before the clinic is scheduled again.

$$(t' - o_i) \cdot c_{i,t,t'} \cdot x_{i,t} \geq f_i \cdot c_{i,t,t'} \cdot x_{i,t} \qquad \forall\, i \in \mathcal{I},\ t, t' \in \mathcal{T} \tag{3.8}$$

To consider the frequency of clinics within the planning horizon, two sets of constraints are utilised. Firstly, we consider the subsets of clinics $\mathcal{I}^{f_i \leq T-t}$ $\forall t \in \mathcal{T}$ that have a maximum frequency less than or equal to the number of remaining days of the planning horizon, from starting day $t \in \mathcal{T}$. These clinics must not be scheduled more than once during f_i days. The following constraints (3.9) ensure that this holds true by requiring the number of times a clinic is scheduled over f_i remaining days is less than or equal to one.

$$\sum_{\tilde{t} \in [t, t+f_i]} x_{i,\tilde{t}} \leq 1 \qquad \forall\, i \in \mathcal{I}^{f_i \leq T-t},\ t \in \mathcal{T} \tag{3.9}$$

Similarly, we consider the sets of clinics $\mathcal{I}^{f_i > T-t}$ $\forall t \in \mathcal{T}$ that have a maximum frequency greater than the remaining length of the planning horizon $\mathcal{T}$ after the considered starting day $t \in \mathcal{T}$. Constraints (3.10) ensure that these clinics are not scheduled more than once during the remaining days of the planning horizon.

$$\sum_{\tilde{t} \in [t, T]} x_{i,\tilde{t}} \leq 1 \qquad \forall\, i \in \mathcal{I}^{f_i > T-t},\ t \in \mathcal{T} \tag{3.10}$$

As discussed in Section 3.1.4, the Welsh Blood Service requires each clinic location to be visited at least once per year. Constraints (3.11) ensure this.

$$(t' - o_i) \cdot c_{i,t,t'} \cdot x_{i,t} \leq 365 \qquad \forall\, i \in \mathcal{I},\ t, t' \in \mathcal{T} \tag{3.11}$$

Weekday and Seasonal Availability Constraints

Some clinics may only be available on specific weekdays, or specific times of the year; thus we have weekday and seasonal availability, ensured by constraints (3.12) and (3.13), respectively.

$$\sum_{t \in \mathcal{T}} c_{i,t,t'} \cdot x_{i,t} \leq a_{i,t'}^{\text{weekday}} \qquad \forall\, i \in \mathcal{I},\ t' \in \mathcal{T} \tag{3.12}$$

$$\sum_{t \in \mathcal{T}} c_{i,t,t'} \cdot x_{i,t} \leq a_{i,t'}^{\text{seasonal}} \qquad \forall\, i \in \mathcal{I},\ t' \in \mathcal{T} \tag{3.13}$$

If for example, a given clinic i is not available on a Monday, this relates to the weekday availability and the $a_{i,t}^{\text{weekday}}$ that corresponds to Monday(s) would be $a_{i,t}^{\text{weekday}} = 0$.

Minimum and Maximum Clinics Per Time

It must be ensured that each clinic is not scheduled more than once on any given day (3.14).

$$\sum_{t \in \mathcal{T}} c_{i,t,t'} \cdot x_{i,t} \leq 1 \qquad \forall\, i \in \mathcal{I},\ t' \in \mathcal{T} \tag{3.14}$$

Each region has a maximum number of clinics that can be scheduled to occur on a given day due to resource constraints. Thus constraints (3.15) prevent this from being exceeded.

$$\sum_{i \in \mathcal{I}^r} \sum_{t \in \mathcal{T}} c_{i,t,t'} \cdot x_{i,t} \leq R_{t,r}^{max} \qquad \forall\, r \in \mathcal{R},\ t' \in \mathcal{T} \tag{3.15}$$

Similarly, each region also has a maximum number of clinic days it can operate over a calendar week due to resource constraints. To ensure these are met, we have the following constraints.

$$\sum_{i \in \mathcal{I}^r} \sum_{t' \in \mathcal{T}_w} \sum_{t \in \mathcal{T}} c_{i,t,t'} \cdot x_{i,t} \leq R^{max}_{w,r} \qquad \forall\, r \in \mathcal{R},\ w \in \mathcal{W} \tag{3.16}$$

Due to workforce contractual hours at the Welsh Blood Service, it must be ensured that the minimum number of clinic days are scheduled per region over the given planning horizon. Thus, we have the following constraints.

$$\sum_{i \in \mathcal{I}^r} \sum_{t' \in \mathcal{T}} \sum_{t \in \mathcal{T}} c_{i,t,t'} \cdot x_{i,t} \geq R^{min}_{\mathcal{T},r} \qquad \forall r \in \mathcal{R} \tag{3.17}$$

Constraints for Conflicting Clinics

For the pairs of clinics that must not be scheduled within a given number of days of each other, we include the following constraints.

$$(t' - o_{i_2}) \cdot c_{i_1,t,t'} \cdot x_{i_1,t} \geq b_{i_1,i_2} \cdot c_{i_1,t,t'} \cdot x_{i_1,t} \qquad \forall\, (i_1, i_2) \in \mathcal{N}_b,\ b \in \mathbb{N},\ t, t' \in \mathcal{T} \tag{3.18}$$

$$(t' - o_{i_1}) \cdot c_{i_2,t,t'} \cdot x_{i_2,t} \geq b_{i_1,i_2} \cdot c_{i_2,t,t'} \cdot x_{i_2,t} \qquad \forall\, (i_1, i_2) \in \mathcal{N}_b,\ b \in \mathbb{N},\ t, t' \in \mathcal{T} \tag{3.19}$$

Constraints (3.18) and (3.19) ensure this considers clinics that were scheduled in previous planning horizons, whilst constraints (3.20) considers only days within the given planning horizon, where $q = min\{b_{i_1,i_2}, T - t\}$ i.e. the smallest of either the minimum number of days between the given conflicting clinics, or the remaining number of days of the planning horizon.

$$\sum_{\tilde{t} \in [t, t+q]} x_{i_1,\tilde{t}} + \sum_{\tilde{t} \in [t, t+q]} x_{i_2,\tilde{t}} \leq \ 1 \quad \forall\, (i_1, i_2) \in \mathcal{N}_b,\ b \in \mathbb{N},\ t \in \mathcal{T} \tag{3.20}$$

The following constraints (3.21) ensure that no public clinics within 5 miles of each other are scheduled on the same day. This is to minimise the risk of donors attending an alternative clinic due to a potential overlap of donor panels between clinics in close proximity to each other.

$$c_{i_1,t,t'} \cdot x_{i_1,t} + c_{i_2,t,t'} \cdot x_{i_2,t} \leq 1 \qquad \forall\ (i_1, i_2) \in \mathcal{Z}_j,\ t, t' \in \mathcal{T} \tag{3.21}$$

Constraints by Venue Type

Due to limited resources, a maximum of two trailer clinics can be scheduled to occur on the same day in the South East region, $r = \text{SE}$. Thus, we include constraints (3.22).

$$\sum_{i \in \mathcal{I}^{\text{SE}} \cup \mathcal{I}^{\text{trailer}}} \sum_{t \in \mathcal{T}} c_{i,t,t'} \cdot x_{i,t} \leq 2 \qquad \forall t' \in \mathcal{T} \tag{3.22}$$

Clinic Tour Constraints

To ensure that clinic tours are scheduled in the correct sequence, we make use of constraints (3.23) where we consider only the days that the tour can run fully; this excludes the last $(K-1)$ days to ensure that there are at least K consecutive days available from the starting day of the first clinic of a tour, for the whole tour to be scheduled.

$$c_{i_1,t,t} \cdot x_{i_1,t} = c_{i_k,t+(k-1),t+(k-1)} \cdot x_{i_k,t+(k-1)} \qquad \forall\ k \in \mathbb{N}_{\leq K},\ (i_1, ..., i_K) \in \mathcal{I}^{\text{tours}}, \quad \forall t \in [t_1, t_{T-(K-1)}] \tag{3.23}$$

Each clinic within a tour must only be scheduled as part of said tour and should not be scheduled separately. The above constraint does not ensure that clinic two or three of a tour are not scheduled on the first $(K-1)$ days. Therefore the following constraints ensure that these cases are considered.

$$c_{i_k,t,t} \cdot x_{i_k,t} = 0 \qquad \forall\ k \in \mathbb{N}_{\leq K} \setminus \{1\},\ (i_2, ..., i_K) \in \mathcal{I}^{\text{tours}},\ t \in [t_1, t_{k-1}] \tag{3.24}$$

The following example demonstrates how both constraints (3.23) and (3.24) work: Consider a clinic tour consisting of three clinics, namely clinic 'A', clinic 'B' and clinic 'C'. In this case, $K = 3$. To ensure that these three clinics are scheduled in the correct sequence, constraint (3.23) requires that any time clinic 'A' is scheduled to start on day $t \in [t_1, t_{T-(K-1)}]$, then clinic 'B' is scheduled to start on day $t+k-1 = t+1$ and likewise, clinic 'C' is scheduled to start on day $t+k-1 = t+2$. These constraints also enable the converse to hold; that if clinic 'A' is not scheduled over the days $t \in [t_1, t_{T-2}]$, then neither is clinic 'B' over the days $t \in [t_2, t_{T-1}]$ or clinic 'C' over the days $t \in [t_3, T]$. Constraints (3.24) restrict clinic 'B' from being scheduled individually on the first day of the planning horizon, and clinic 'C' from being scheduled individually during the first two days of the planning horizon.

Workforce Constraints

Clinic-based workers require a day off following a 12 hour working day. This needs to be considered at the clinic scheduling level, especially for the three regions (namely South West, North East and North West) that only have enough workers to form one clinic team, as this means no clinics can be scheduled on a day following a clinic that exceeds 12 working hours. Constraints (3.25) ensure that this is considered, where $i' \in \mathcal{I}^r \cup \mathcal{L}_{12}$ denotes a clinic in region $r \in \mathcal{R}$ that has a length of day greater or equal to 12 hours. The first day of the planning horizon is excluded from the summation over all $\tilde{t}$ since these constraints consider only days following long working days, and thus this cannot be the case for the first working day of the horizon.

$$c_{i',t,t'} \cdot x_{i',t'} + \sum_{i \in \mathcal{I}^r} \sum_{\tilde{t} \in \mathcal{T} \setminus \{t_1\}} c_{i,\tilde{t},t'+1} \cdot x_{i,\tilde{t}} \leq R^{max}_{t,r} \qquad \forall\, i' \in \mathcal{I}^r \cup \mathcal{L}_{12},\ r \in \mathcal{R},$$

$$\forall t, t' \in \mathcal{T} \setminus \{T\} \quad (3.25)$$

Constraints (3.25) work by ensuring that the sum of all clinics in a given region scheduled on the day following a clinic that exceeds 12 working hours is one less than the maximum per day, by including the addition of the 12-hour clinic the day prior.

Clinic workers also require a day off before they work a clinic tour. Since workers from the South East region are more likely to work a clinic tour, we need to ensure that at most, the number of clinics scheduled on the day before a clinic tour is due to begin, is one less than the maximum per day per region i.e. $R_{t,r}^{max} - 1$. Therefore, formulated in a similar manner to constraints (3.25), we have the following constraints to ensure this is the case.

$$c_{i_1,t,t} \cdot x_{i_1,t} + \sum_{i \in \mathcal{I}^r} \sum_{\tilde{t} \in \mathcal{T} \setminus \{T\}} c_{i,\hat{t},t-1} \cdot x_{i,\tilde{t}} \leq R_{t,r}^{max} \qquad \forall\, i_1 \in \mathcal{I}^r \cup \mathcal{I}^{\text{tours}},\ r \in \mathcal{R}, \quad \forall t \in \mathcal{T} \setminus \{t_1\} \tag{3.26}$$

Constraints for Obligatory Clinics

The clinic that takes place at the WBS headquarters must be scheduled to occur on a regular basis. In addition, we utilise dummy clinics for annual leave and training of clinic-based workers, and these are to be scheduled to occur every weekday, for employees to be assigned to a corresponding dummy clinic when they have booked annual leave or training. The following constraints (3.27) ensure that all obligatory clinics are scheduled as required, where g_i denotes the number of days that clinic $i \in \mathcal{I}$ must run for each time it is scheduled.

$$\sum_{t \in \mathcal{T}} \sum_{t' \in \mathcal{T}} c_{i,t,t'} \cdot x_{i,t} = g_i \qquad \forall i \in \mathcal{I}^{\text{forced}} \tag{3.27}$$

Decision Variables and (their) Domains

The following constraints apply to the decision variables to ensure that they are non-negative, and are binary in the case of the variable $x_{i,t}$.

$$x_{i,t} \in \{0,1\} \qquad \forall\, i \in \mathcal{I},\ t \in \mathcal{T} \tag{3.28}$$

$$\Delta_w^+ \in \mathbb{R}_{\geq 0} \qquad \forall t_w \in \mathcal{T}, w \in \mathcal{W} \tag{3.29}$$

3.3.4 Summary

In this chapter, the current practice of clinic scheduling at the Welsh Blood Service has been described, including the format of their blood donation clinics and the limitations of the current practice. A formulation of the Blood Donation Clinic Scheduling Problem is presented in the form of a linear programme to schedule clinics optimally over a given planning horizon. All parameters have been introduced, along with three alternative objective functions; minimisation of the number of clinic days scheduled, minimisation of overcollection, and minimisation of both the number of clinic days scheduled and overcollection. All constraints have been formulated to ensure that the output of the model is a feasible and realistic solution for the WBS.

This model addresses research question one – ‘how can mathematical modelling help to schedule the WBS clinics more efficiently?’ – by reducing monetary costs associated with scheduling more clinics than is necessary to meet demand and/or reducing potential wastage of blood product collected in excess of demand. Additionally, the model aims to streamline the clinic scheduling process and decrease the time taken to generate an initial four-week clinic schedule. This model will be combined with the model presented in Chapter 4 to construct a decision support tool for the Welsh Blood Service, with results of the BDCSM presented in Chapter 5.

In the next chapter, Chapter 4, the Blood Collection Workforce Scheduling Problem is presented following a description of the current practice at the WBS for scheduling clinic-based workers.

Chapter 4

A Blood Collection Workforce Scheduling Problem

This chapter describes the current practice for scheduling the blood collection workforce at the Welsh Blood Service (WBS), the limitations of this practice, and proposes a mathematical model to improve the efficiency of the clinic-based workforce scheduling process. This chapter is structured as follows; Section 4.1 describes the current practice of scheduling the clinic-based workforce at the WBS, while Section 4.2 presents the mathematical formulation of the Blood Collection Workforce Scheduling Problem (BCWSP) with all relevant parameters. The Blood Collection Workforce Scheduling Model (BCWSM) formulation is presented in Section 4.3 with a description of all decision variables, objective functions and constraints. A modified version of the BCWSM is presented in the subsequent Section 4.4 to improve upon solutions of the original model.

4.1 Current Practice

At present, the Welsh Blood Service (WBS) organises their clinic workforce into set teams with limited flexibility between these teams. Each worker is assigned to one of four regions, based on their home address and willingness to travel. The vast majority of the time, each worker will only be assigned to clinics in the corresponding region, unless there is a shortage of staff; in this case, workers from other regions that are willing to cover another region may be asked to do so. A team is assigned to each scheduled clinic (within the same region) far in advance, with a more detailed workforce schedule created closer to the time (usually four to five weeks prior to the first week of a schedule) considering individual workers shifts, annual leave, training etc. Similarly to clinic schedules determined in Chapter 3, workers' schedules are planned over four-week periods at a time.

4.1.1 Clinic Teams

One-Team Regions

In the three smaller regions, namely South West, North West and North East, there is only enough clinic-based workers to form one clinic team. This is due to the smaller population in these regions, with at most one clinic scheduled per day in each of these regions. Some members of the workforce in one of the north regions may be willing to work clinics in the alternative north region if there is a staff shortage, but due to the distance between team bases, this only occurs in rare circumstances.

Multi-Team Region

The South East region consists of four teams of clinic-based workers with little flexibility between these teams. Two of these teams are assigned to only community clinics (at a venue) whilst the other two teams are assigned to only trailer clinics (bloodmobile). In cases where one team may have a staff shortage on a given day, workers from another team that are willing to do so will join the team that requires cover. Up to four clinics per day can be scheduled in this region, with the population density high enough to support this.

For all regions, if there are not any workers able to cover absences, then the clinic will operate at a reduced capacity by decreasing the number of donor chairs open and limiting the number of 'walk-in' appointments on the day proportionate to the staff shortage and capacity of the clinic.

4.1.2 Roles and Skills

Clinic Roles

There are a variety of roles required at blood donation clinics to ensure effective collection of blood.

- **Registered Nurse (RN):** At least one nurse is required to be present at each clinic to assist with eligibility screening of donors to deal with more complicated elements such as diseases or medications, etc.
- **Supervisor:** Each clinic requires a Clinic Supervisor i.e. the manager of the clinic for its duration. This role may be assigned to either a qualified Supervisor or a Deputy Supervisor if there is not an available Supervisor within the same region.
- **Clinic Care Assistant (CCA):** Multiple CCAs are required to work each scheduled clinic, with the specific number required dependant on the capacity of the clinic. A CCA is (in most cases) a trained venepuncturist and therefore able to work at any station within a clinic.
- **Clinic Service Assistant (CSA):** A CSA is similar to a CCA with the main difference being that CSAs are not venepuncturists, and therefore may assist with donors at all stages of the donation process except for venepuncture.

Each clinic has a minimum requirement for each clinic role (with CCA and CSA considered together) and these numbers are dependant on the capacity of the clinic. For small clinics, only one nurse may be necessary whereas for much larger clinics, having two nurses can help to drive efficiency and reduce bottlenecks in the eligibility screening process. The official Welsh Blood Service guidance on the clinic role configuration is displayed in Table 4.1, which breaks down the specific roles required at a clinic. Clinic workers that are CCAs and CSAs are assigned to these more specific roles, with the exclusion of the roles of Supervisor and RN. However, the 'pod care' activity requires venepuncture and therefore only CCAs can be assigned to this role.

Driving Roles

Each whole blood clinic (with the exception of the clinic located at the WBS Headquarters in Talbot Green) has vehicle requirements in order to operate. We describe the function of each vehicle type below.

- **Lorry:** Utilised to transport clinic equipment including donation chairs, privacy screens and blood bags. They require unloading and loading at the beginning and end respectively of each day in use. Lorries are necessary for any community clinics as they take place in a booked venue, and all clinic equipment needs to be transported to and from the clinic site.
- **Minibus:** These are ideally utilised for every clinic that takes place outside of WBS premises to transport all clinic workers from the region base to the clinic and back at the end of the working day.
- **Support Vehicle:** These are essentially a 'bloodmobile' i.e. a vehicle that acts as a clinic when it is parked at a designated trailer clinic site. All clinic equipment is already in place and therefore no unloading or loading is necessary. Each support vehicle has a capacity of six donation chairs.
- **Mobile Donation Unit:** Similarly to a support vehicle, a mobile donation unit acts as a clinic when it is parked at a designated trailer site. The difference between the two vehicle types is that a mobile donation unit has a smaller capacity of only three donation chairs.

Table 4.1: Clinic Role Configurations per Clinic Size

<table>
<tr><th></th><th></th><th colspan="9">Number of Workers Advised per Activity/ Role</th></tr>
<tr><th>Clinic Size (Number of Chairs)</th><th>Clinic Type</th><th>Registration</th><th>Screening</th><th>Pod Care</th><th>Pod Support</th><th>Post-Donation Care</th><th>Quality Check</th><th>Supervisor</th><th>Registered Nurse</th><th>Total</th></tr>
<tr><td>3</td><td>Trailer</td><td colspan="2">1</td><td colspan="4">1</td><td>1</td><td>1</td><td>4</td></tr>
<tr><td>3</td><td>Community</td><td colspan="2">1</td><td colspan="2">1</td><td colspan="2">1</td><td>1</td><td>1</td><td>5</td></tr>
<tr><td>4</td><td>Community</td><td>1</td><td>1</td><td colspan="2">2</td><td colspan="2">1</td><td>1</td><td>1</td><td>7</td></tr>
<tr><td>5</td><td>Community</td><td>1</td><td>2</td><td colspan="2">2</td><td colspan="2">1</td><td>1</td><td>1</td><td>8</td></tr>
<tr><td>6</td><td>Trailer</td><td>1</td><td>2</td><td>2</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td><td>10</td></tr>
<tr><td>6</td><td>Community</td><td>1</td><td>2</td><td>2</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td><td>10</td></tr>
<tr><td>7</td><td>Community</td><td>1</td><td>2</td><td>2</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td><td>10</td></tr>
<tr><td>8</td><td>Community</td><td>1</td><td>3</td><td>2</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td><td>11</td></tr>
<tr><td>9</td><td>Community</td><td>1</td><td>3</td><td>3</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td><td>12</td></tr>
<tr><td>10</td><td>Community</td><td>1</td><td>3</td><td>3</td><td>1</td><td>1</td><td>1</td><td>1</td><td>2</td><td>13</td></tr>
<tr><td>11</td><td>Community</td><td>2</td><td>4</td><td>4</td><td>1</td><td>1</td><td colspan="2">1</td><td>2</td><td>15</td></tr>
<tr><td>12</td><td>Community</td><td>2</td><td>4</td><td>4</td><td>1</td><td>1</td><td>1</td><td>1</td><td>2</td><td>16</td></tr>
</table>

All activities excluding Registered Nurse (RN) and Supervisor are reserved for all other clinic workers i.e. Clinic Care Assistants (CCAs) and Clinic Support Assistants (CSAs). A Supervisor role may be filled by a Deputy Supervisor, otherwise a Deputy Supervisor is assigned as a CCA. For example, a four-chair community clinic requires a minimum of one Supervisor, one RN and two CCAs/CSAs.

Each community clinic requires both a lorry and a minibus. Trailer clinics only occur in the South East and North West regions, and each of these clinics require either a support vehicle (for South East clinics) or a mobile donation unit (for North West clinics), in addition to a minibus. If there are not enough certified minibus drivers available to cover all demand within a region on a given day, clinic-based workers are asked to make their own way to their assigned clinic, for all clinics without a minibus.

Only a select number of workers in each region are qualified and insured to drive specific vehicles. Generally, if a worker is trained to drive one vehicle type, then they are trained to drive all other vehicle types required in their region e.g. a worker in the South East region that is qualified to drive a lorry with likely also be qualified to drive a minibus and a support vehicle. However, there are a few exceptions with some drivers only being trained and/or willing to drive a limited number of vehicle types.

At present, the number of trained drivers in some regions seems restrictive, which is likely to cause issues meeting demand with absences due to sickness and/or annual leave amongst the drivers.

4.1.3 Contracted Hours and Agreements

Working Day Patterns

For the majority of clinic workers, their agreed working days are Monday to Friday with the occasional Saturday clinic. However, some workers have more specific working day patterns due to childcare issues or an agreed work/life balance. For example, some workers may have a specific weekday that they do not work, while others may have a more complicated pattern such as differing specific weekdays on alternating months.

Contracted Hours

There is a range of 12 different weekly contracted hour amounts across the clinic-based workforce, with the minimum being 18 hours and maximum being 37.5 hours. Full-time is considered to be 37.5 hours per week, with anything less than this deemed part-time. Clinic workers are paid for their contracted hours regardless of if all hours were actually worked, due to the variability of demand and clinic schedules.

Annual Leave

Every worker is entitled to annual leave, with the number of days per year dependant on the length of time a given worker has been employed by the NHS. This ranges from 27 days (plus Bank Holidays) for under five years of NHS employment, 29 days (plus Bank Holidays) for those with 5-10 years of NHS employment, and 33 days (plus Bank Holidays) for those employed with the NHS for over ten years.

Training

Each clinic worker is required to complete a minimum of three days mandatory training per year. Each training day is a standard working day of 7.5 hours. Most training is regarding health and safety such as manual handling for clinic equipment, and needs to be completed each year as certification expires.

Salary

Salary for clinic workers follow the standard NHS Wales model and is determined by both employee skill band and the duration of their employment. Different clinic roles, as described above, have a different salary band, as displayed in Table 4.2 retrieved from online data [66]. Each band has a corresponding salary range, with those with less than one year experience in the role starting at the lower bound of the range. For each year that an employee gains experience in the role, their salary will increase, until the upper bound of the range is reached; This is six years for bands two to four, seven years for band five and eight years for band six. The salary is based on full-time work i.e. 37.5 hours per week, and is calculated pro-rata for any part-time employees. We have calculated the average hourly pay rate per salary band (see Table 4.2) by taking the mid-point of the salary range per band and dividing by 52 weeks, and then dividing again by 37.5 hours.

Table 4.2: NHS Salary Bands as of April 2020

Clinic Role	**Salary Band**	**Salary Range**	**Average Hourly Rate**
CSA	2	£18,005 - £19,337	£9.58
CCA	3	£19,737 - £21,142	£10.48
Deputy Supervisor	4	£21,892 - £24,157	£11.81
Supervisor	5	£24,904 - £30,615	£14.24
RN	5	£24,904 - £30,615	£14.24
Senior RN	6	£31,365 - £37,890	£17.76

Overtime

Overtime is voluntary and assigned on a 'first come, first serve' basis. Some employees prefer not to work any overtime, whilst others are more flexible and are even willing to work overtime in a different region from their own base region.

Overtime is defined as any worked hours in excess of total contracted hours over a four-week period, e.g. if a full-time worker (with 37.5 contracted hours per week) works 40 hours one week, but 36 hours each of the remaining weeks of the four-week period, this is a total of 148 hours and is less than their contracted 150 hours, and therefore no overtime was worked.

It is each employee's decision if they would prefer overtime to be paid or be awarded time off in lieu. The vast majority of the WBS clinic workers are paid overtime, with the rate of pay for overtime hours depends on the total hours worked over the four-week period. If the total number of worked hours is below the full-time amount (150 hours), then the hourly rate is the same rate as usual for that worker. If the total worked hours over the four-week period exceeds full-time, then the hourly pay rate for each hour in excess of this is at 1.5 times the usual rate for that worker. Therefore, it is in the interest of WBS to minimise these costs, particularly for overtime at the increased rate.

4.1.4 Limitations of Current Practice

The current practice has a range of limitations on efficiency that may cause unnecessary costs to the organisation. Firstly, the lack of flexibility in the workforce based in the South East region drives inefficiency, as the team size is fixed (often the size necessary to operate the largest clinics) but the clinic sizes and staffing requirements vary considerably. This model results in the WBS requiring a higher level of workers in the South East region than is likely necessary if a more flexible staffing model were introduced. Usually, aside from any workers that have booked annual leave or training, or are absent due to sickness, all members of a clinic team will be assigned to a clinic, regardless of if the clinic requires that number of workers. This often leads to clinics being overstaffed in order to fulfil contractual hours of workers. Occasionally, an overstaffed clinic may provide opportunity to open an additional donation chair, and perhaps accept more ‘walk-in’ appointments. However, this is often prohibited due to the physical capacity of the clinic venue and related health and safety regulations.

Secondly, the Welsh Blood Service uses a considerable amount of overtime hours amongst their clinic-based workers to enable all scheduled clinics to run effectively. This implies that there is an insufficient number of workers and/or contracted hours to adequately manage typical clinic schedules. However, with the Blood Donation Clinic Scheduling Model (BDCSM) presented in Chapter 3 aiming to schedule fewer clinics to meet demand, the resulting clinic schedules will likely require less overtime.

Additionally, the optimal levels of staffing and balance of skills and roles is unknown by the organisation. Currently, if an employee leaves the WBS, their same role (including number of contractual hours) is usually advertised to be filled. The staffing level decisions are not informed by data, but by ensuring each team has enough workers to operate the largest capacity clinic in the corresponding region with absences (annual leave and sickness) considered.

Finally, the workforce planning team typically spend an average of 48 hours in total to produce a four-week schedule for clinic-based workers for all regions. This includes inputting data such as annual leave into a digital rostering system, but the system is not automated and thus schedules are required to be created manually. This process is extremely time-consuming for the planning team and would benefit from an increase of efficiency.

We aim to address these limitations in the following model formulation, with the main objective to reduce costs through minimisation of overtime pay and to streamline the workforce scheduling process for clinic-based workers, whilst considering all of the required constraints. In addition, the workers in the South East region are considered in the model collectively as one pool of staff in replacement of four predetermined teams to promote better use of contracted hours.

4.2 Problem Formulation

The Blood Collection Workforce Scheduling Model forms the second stage of the scheduling model and assigns clinic-based workers to each clinic scheduled in the first stage of the model, the Blood Donation Clinic Scheduling Model (BDCSM).

4.2.1 Planning Horizon

Let $\mathcal{T} = \{1, \ldots, T\}$ be the set of discrete times, within the planning horizon. They are equally incremented as days, and thus we will relate to each time increment as one day, $t \in \mathcal{T}$. Each planning horizon is to begin on a Monday, with each full calendar week during the planning horizon within the set $\mathcal{W} = \{1, \ldots, T/7\}$ of weeks.

4.2.2 Workforce

Set $\mathcal{S}$ depicts the set of all clinic-based workers, where $s \in \mathcal{S}$ represents a unique worker. Set $\mathcal{J}$ depicts the set of all potential roles of clinic workers, where $\mathcal{S}_j \subset \mathcal{S}$ represents a subset of workers who can perform role j. We also need to consider the following for each worker:

- **Workforce Availability:** Let $a_{s,t}$ denote the availability of worker $s \in \mathcal{S}$ such that;

$$a_{s,t} = \begin{cases} 1 & \text{if worker } s \in \mathcal{S} \text{ is available to work at a clinic on day } t \in \mathcal{T} \\ 0 & \text{otherwise} \end{cases} \tag{4.1}$$

  This availability considers any booked annual leave, pre-scheduled training days and a worker's agreed working day pattern.

- **Annual Leave:** Let $g^{\text{leave}}_{s,t}$ denote the booked annual leave of worker $s \in \mathcal{S}$ such that;

$$g^{\text{leave}}_{s,t} = \begin{cases} 1 & \text{if worker } s \in \mathcal{S} \text{ is to have annual leave on day } t \in \mathcal{T} \\ 0 & \text{otherwise} \end{cases} \tag{4.2}$$

- **Training:** Let $g^{\text{training}}_{s,t}$ denote the pre-determined days of training for worker $s \in \mathcal{S}$ such that;

$$g^{\text{training}}_{s,t} = \begin{cases} 1 & \text{if worker } s \in \mathcal{S} \text{ is to have training on day } t \in \mathcal{T} \\ 0 & \text{otherwise} \end{cases} \tag{4.3}$$

- **Contracted Hours:** Let h_s denote the number of contracted hours per week for worker $s \in \mathcal{S}$. Let C denote the number of contracted hours per week for a full-time worker.

- **Region:** Let $\mathcal{R}$ denote the set of regions. Let s_r denote a worker $s \in \mathcal{S}$ that is based in the region $r \in \mathcal{R}$.

- **Preferences:** Each worker $s \in \mathcal{S}$ may provide preferences regarding what types of clinics they are to be assigned to work. Let p_s^{overtime} denote whether worker s is willing to work overtime, such that;

$$p_s^{\text{overtime}} = \begin{cases} 1 & \text{if worker } s \in \mathcal{S} \text{ is unable to be assigned to work overtime} \\ 0 & \text{otherwise} \end{cases} \tag{4.4}$$

  Let p_s^{tour} denote whether worker s is willing to work at trailer clinic, such that;

$$p_s^{\text{tour}} = \begin{cases} 1 & \text{if worker } s \in \mathcal{S} \text{ is unable to be assigned to a tour clinic} \\ 0 & \text{otherwise} \end{cases} \tag{4.5}$$

- **Overtime Pay:** For simplification, we assume that all workers are paid for overtime, since only one worker currently receives time off in lieu instead of additional pay and this is unlikely to have an impact on overall annual leave. As described in Section 4.1.3, there are two pay rates for overtime: a base hourly rate which is the corresponding estimated hourly pay for a given worker, and an increased rate of 1.5 times the base rate. The base rate is paid for all hours worked up to full-time hours over the four-week period (150 hours), while the increased rate is for any hours worked in excess of this. Let b_s denote the base hourly rate of worker s, displayed as average hourly rate in Table 4.2. Therefore the increased hourly rate for worker s is given by 1.5 b_s.

4.2.3 Clinics

Set $\mathcal{I}$ depicts a set of clinic locations, with each unique clinic location denoted as i. We use a pre-determined clinic schedule (the output of stage one of the model i.e. the BDCSM) and utilise the notation of the stage one formulation (Chapter 3) such that;

$$x_{i,t} = \begin{cases} 1, & \text{if clinic } i \in \mathcal{I} \text{ is scheduled on day } t \in \mathcal{T}. \\ 0, & \text{otherwise.} \end{cases} \tag{4.6}$$

Each clinic $i \in \mathcal{I}$ has the following attributes.

- **Region:** Let $\mathcal{R}$ denote the set of regions. Let i_r denote the region r that clinic i is located within.

- **Length of Day:** Let $l_{i,t}$ denote the expected length of working day (in hours) of clinic $i \in \mathcal{I}$ on day $t \in \mathcal{T}$. Let $\mathcal{L}_{12}$ be the set of clinics i with length of day greater than or equal to 12 hours i.e. $l_{i,t} \geq 12$.

- **Venue Type:** Let $i^{\text{trailer}} \in \mathcal{I}$ denote the clinics in a trailer venue type and $i^{\text{community}} \in \mathcal{I}$ denote a community venue clinic.

- **Clinic Type:** Let $\mathcal{I}^{\text{tours}} \subset \mathcal{I}$ denote the set of tuples of clinics, $(i_1, ...i_K) \in \mathcal{I}^{\text{tours}}$ where each tuple forms a clinic tour. Let $k \in \mathbb{N}_{\leq K}$ represent each individual day of a tour, where $K \in \mathbb{N}$ denotes the final day number of a tour. Thus, K also denotes the length of a tour i.e. the number of consecutive days it requires each time it is scheduled.

- **Resource Demand:** Let $d_{i,j}$ denote the minimum number of workers required to perform role $j \in \mathcal{J}$ each day that clinic i is open $\forall i \in \mathcal{I}$.

Dummy Clinics

In order to include worker hours for training and annual leave for the workforce, we utilise dummy clinics. Each worker $s \in \mathcal{S}$ has an associated annual leave dummy clinic, $i_s^{\text{leave}} \in \mathcal{I}$, with length of day, l_i equivalent to one day of annual leave based on the contracted hours of worker s. Similarly, there is one dummy training clinic (as the length of training days are the same for all workers) which we denote as $i^{\text{training}} \in \mathcal{I}$. Therefore, let $\mathcal{I}^{\text{standard}} \subset \mathcal{I}$ denote the set of clinics, i, that are actual clinics of the organisation and not dummy clinics.

4.2.4 Working Modes

In order to account for situations where a worker may have more than one role to carry out on a given day (and at a given clinic) i.e. a clinic role and a driving role, we include the concept of working modes. Let $m \in \mathcal{M}_s$ denote a working mode of worker $s \in \mathcal{S}$. Due to all RNs, Supervisors, CCAs and CSAs only having one possible clinic role, there is a maximum of five possible working modes for each of these workers; these consist of the clinic role alone with no driving role, and each driving role in turn alongside the clinic role, as displayed in Table 4.3. However, for those with the job title of Deputy Supervisor, they have two possible clinic roles - Supervisor and CCA. Therefore, the maximum number of modes for a Deputy Supervisor is ten, as displayed in Table 4.4. Since not all clinic workers are willing and/or qualified to have any driving roles, most workers have only one possible working mode.

Table 4.3: Working Modes Example for RN

	Driving Role				**Clinic Role**		
Mode	**Lorry**	**Minibus**	**SV**	**MDU**	**RN**	**Supervisor**	**CCA/CSA**
1	0	0	0	0	1	0	0
2	1	0	0	0	1	0	0
3	0	1	0	0	1	0	0
4	0	0	1	0	1	0	0
5	0	0	0	1	1	0	0

Table 4.4: Working Modes Example for Deputy

	Driving Role				Clinic Role		
Mode	**Lorry**	**Minibus**	**SV**	**MDU**	**RN**	**Supervisor**	**CCA/CSA**
1	0	0	0	0	0	1	0
2	1	0	0	0	0	1	0
3	0	1	0	0	0	1	0
4	0	0	1	0	0	1	0
5	0	0	0	1	0	1	0
6	0	0	0	0	0	0	1
7	1	0	0	0	0	0	1
8	0	1	0	0	0	0	1
9	0	0	1	0	0	0	1
10	0	0	0	1	0	0	1

4.3 Model Formulation

With all parameters introduced, we now introduce the decision variables and their domains as well as the constraints.

4.3.1 Decision Variables

We introduce the binary decision variable $y_{s,m,i,t}$ such that;

$$y_{s,m,i,t} = \begin{cases} 1, & \text{if worker } s \in \mathcal{S} \text{ is assigned to work mode } m \in \mathcal{M}_s \\ & \text{at clinic } i \in \mathcal{I} \text{ on day } t \in \mathcal{T} \\ 0, & \text{otherwise} \end{cases} \tag{4.7}$$

We introduce the following continuous decision variables $z_{s,w}^{\text{over}}$ and $z_{s,w}^{\text{under}}$ to represent weekly overtime and 'undertime' respectively, per worker, s, such that;

$$z_{s,w}^{\text{over}} \in \mathbb{R}_{\geq 0} \tag{4.8}$$

$$z_{s,w}^{\text{under}} \in \mathbb{R}_{\geq 0} \tag{4.9}$$

Table 4.5: Notation and list of abbreviations

Sets	
$\mathcal{I}$	Set of clinics
$\mathcal{I}^{\text{standard}} \subset \mathcal{I}$	Set of standard clinics (not dummy clinics)
$\mathcal{I}^{tours} \subset \mathcal{I}$	Set of clinic tours
$\mathcal{J}$	Set of all potential roles of workers
$\mathcal{K}$	Set of days of a clinic tour
$\mathcal{L}_{12}$	Set of clinics i with $l_{i,t} \geq 12$
$\mathcal{M}_s$	Set of modes for worker $s \in \mathcal{S}$
$\mathcal{M}_{s,j} \subset \mathcal{M}$	Set of modes for worker $s \in \mathcal{S}$ for role $j \in \mathcal{J}$
$\mathcal{M}_{s,\text{sup}} \subset \mathcal{M}$	Set of modes for worker $s \in \mathcal{S}$ with clinic role of Supervisor
$\mathcal{R}$	Set of regions
$\mathcal{S}$	Set of workers
$\mathcal{S}_r$	Set of workers based in region $r \in \mathcal{R}$
$\mathcal{T}$	Set of days
$\mathcal{T}_w$	Set of days in week $w \in \mathcal{W}$
$\mathcal{W}$	Set of weeks
Parameters	
$a_{s,t}$	Availability of worker $s \in \mathcal{S}$ on day $t \in \mathcal{T}$
b_s	Hourly pay for worker $s \in \mathcal{S}$
$\beta_n \in \mathbb{R}$	Objective function weight parameter for the n^{th} term, where $n = \{1, 2, 3\}$
$C \in \mathbb{R}_{\geq 0}$	Full-time weekly contract hours
$d_{i,j}$	Demand for role $j \in \mathcal{J}$ at clinic $i \in \mathcal{I}$
$g_{s,t}^{\text{leave}}$	Annual leave indicator for worker $s \in \mathcal{S}$ on day $t \in \mathcal{T}$
$g_{s,t}^{\text{training}}$	Training indicator for worker $s \in \mathcal{S}$ on day $t \in \mathcal{T}$
h_s	Contracted hours per week for worker $s \in \mathcal{S}$
$i_r \in \mathcal{I}$	Clinic in region $r \in \mathcal{R}$
$i^{\text{community}} \in \mathcal{I}$	Clinic of community venue type
$i_s^{\text{leave}} \in \mathcal{I}$	Dummy clinic for annual leave for worker $s \in \mathcal{S}$
$i^{\text{trailer}} \in \mathcal{I}$	Clinic of trailer venue type
$i^{\text{training}} \in \mathcal{I}$	Dummy training clinic
$(i_1, ..., i_K) \in \mathcal{I}^{\text{tours}}$	Clinic tour consisting of $K \in \mathbb{N}$ clinics
$l_{i,t}$	Working day duration for day $t \in \mathcal{T}$ of clinic $i \in \mathcal{I}$
p_s^{overtime}	Overtime preference indicator for worker $s \in \mathcal{S}$
p_s^{tour}	Tour preference indicator for worker $s \in \mathcal{S}$
$x_{i,t}$	Schedule indicator for clinic i on day $t \in \mathcal{T}$
Decision Variables	
$v_{1,s}$	Number of hours worker $s \in \mathcal{S}$ is scheduled in excess of full time
$v_{2,s}$	Number of overtime hours worker $s \in \mathcal{S}$ is scheduled
$y_{s,m,i,t}$	1 if worker $s \in \mathcal{S}$ is scheduled to work clinic $i \in \mathcal{I}$ in mode $m \in \mathcal{M}_s$ on day $t \in \mathcal{T}$, 0 otherwise
$z_{s,w}^{\text{over}} \in \mathbb{R}$	Scheduled overtime for worker $s \in \mathcal{S}$ for week $w \in \mathcal{W}$
$z_{s,w}^{\text{under}} \in \mathbb{R}$	Scheduled 'undertime' for worker $s \in \mathcal{S}$ for week $w \in \mathcal{W}$

We make use of variables $v_{1,s}$ and $v_{2,s}$ to indicate the rate of overtime pay for each worker, $s \in \mathcal{S}$ such that;

$$v_{1,s} = \begin{cases} (\sum_{m \in \mathcal{M}_s} \sum_{i \in \mathcal{I}} \sum_{t \in \mathcal{T}} y_{s,m,i,t} \cdot l_{i,t}) - 37.5 \cdot |\mathcal{W}|, & \text{if } \sum_{m \in \mathcal{M}_s} \sum_{i \in \mathcal{I}} \sum_{t \in \mathcal{T}} y_{s,m,i,t} \cdot l_{i,t} > \\ & \quad 37.5 \cdot |\mathcal{W}| \\ & \text{for worker } s \in \mathcal{S}. \\ 0, & \text{otherwise.} \end{cases} \tag{4.10}$$

$$v_{2,s} = \begin{cases} (\sum_{m \in \mathcal{M}_s} \sum_{i \in \mathcal{I}} \sum_{t \in \mathcal{T}} y_{s,m,i,t} \cdot l_{i,t}) - h_s, & \text{if } \sum_{m \in \mathcal{M}_s} \sum_{i \in \mathcal{I}} \sum_{t \in \mathcal{T}} y_{s,m,i,t} \cdot l_{i,t} > h_s \\ & \text{for worker } s \in \mathcal{S}. \\ 0, & \text{otherwise.} \end{cases} \tag{4.11}$$

Thus, $v_{1,s}$ is the number of hours that worker s is scheduled for overtime that exceeds full-time hours and will be paid at the increased rate of 1.5 times their usual hourly pay. Similarly, $v_{2,s}$ is the number of hours that worker, s, is scheduled for overtime in total. This allows us to calculate the correct payment for scheduled overtime hours for each worker in the following objective functions.

4.3.2 Objective Function

We consider two alternative objective functions; one to minimise the total cost of scheduled overtime and one to minimise the total cost of all scheduled hours. The former relates directly to the priority of the WBS to reduce overtime where possible, while the latter is motivated by the possibility of further reducing costs for the WBS. We formulate the two alternative objective functions as follows:

Objective Function One

$$\text{Minimise} \sum_{s \in \mathcal{S}} \left(b_s \cdot v_{2,s} \sum_{m \in \mathcal{M}_s} \sum_{i \in \mathcal{I}} \sum_{t \in \mathcal{T}} y_{s,m,i,t} \cdot l_{i,t} + \frac{b_s}{2} \cdot v_{1,s} \sum_{m \in \mathcal{M}_s} \sum_{i \in \mathcal{I}} \sum_{t \in \mathcal{T}} y_{s,m,i,t} \cdot l_{i,t} \right) + \beta \sum_{s \in \mathcal{S}_{\text{dep}}} \left(\sum_{m \in \mathcal{M}_{\text{sup}}} \sum_{i \in \mathcal{I}} \sum_{t \in \mathcal{T}} y_{s,m,i,t} \right) \quad (4.12)$$

Objective Function Two

$$\text{Minimise} \sum_{s \in \mathcal{S}} \left(b_s \sum_{m \in \mathcal{M}_s} \sum_{i \in \mathcal{I}} \sum_{t \in \mathcal{T}} y_{s,m,i,t} \cdot l_{i,t} + \frac{b_s}{2} \cdot v_{1,s} \sum_{m \in \mathcal{M}_s} \sum_{i \in \mathcal{I}} \sum_{t \in \mathcal{T}} y_{s,m,i,t} \cdot l_{i,t} \right) + \beta \sum_{s \in \mathcal{S}_{\text{dep}}} \left(\sum_{m \in \mathcal{M}_{\text{sup}}} \sum_{i \in \mathcal{I}} \sum_{t \in \mathcal{T}} y_{s,m,i,t} \right) \quad (4.13)$$

We include a penalty function (the second term in each objective function) where the constant β is some constant to enable the penalty value to be comparable to the values of other terms in the objective function and therefore considered in the optimisation process. We utilise this penalty function to discourage workers with the title of Deputy Supervisor being scheduled to fulfil the clinic role of Supervisor for a given shift if other Supervisors are available to work this shift.

The inclusion of variables $v_{1,s}$ and $v_{2,s}$ enables us to consider the costs as required. As shown in 4.10, the variable $v_{1,s}$ represents the difference between the total scheduled hours for worker $s \in \mathcal{S}$ over the four-week planning horizon and the threshold for overtime paid at the increased rate, if the former is greater than the latter, i.e. $v_{1,s}$ gives the number of hours worker s is scheduled to work overtime at the increased pay rate. Similarly, as shown in 4.11, $v_{2,s}$ represents the difference between the total scheduled hours and the total number of contracted hours, h_s over the four-week planning horizon for worker $s \in \mathcal{S}$, i.e. $v_{2,s}$ gives the number of hours of overtime scheduled for worker, s.

Thus, objective function one (4.12) aims to minimise the total cost of overtime over the planning horizon and to minimise the number of Deputy Supervisors assigned to the role of Supervisor, whilst objective function two (4.13) aims to minimise the total cost of all scheduled hours and again to minimise the number of Deputy Supervisors assigned to the role of Supervisor.

4.3.3 Constraints

Resource Demand Constraints

Each standard clinic, $i \in \mathcal{I}^{\text{standard}}$, on any day $t \in \mathcal{T}$ that it is scheduled, must have enough workers assigned to the clinic to meet the clinic's resource demands, $d_{i,j}$. This is to ensure that the clinic can operate effectively, with the predetermined number of workers per role (as displayed in Table 4.1) to be fulfilled.

$$\sum_{s \in \mathcal{S}} \sum_{m \in \mathcal{M}_{s,j}} y_{s,m,i,t} \geq x_{i,t} \cdot d_{i,j} \qquad \forall i \in \mathcal{I}^{\text{standard}},\ t \in \mathcal{T},\ j \in \mathcal{J} \tag{4.14}$$

These constraints ensure that for every role $j \in \mathcal{J}$ that is required to occur at clinic $i \in \mathcal{I}$, the minimum number of workers per role are scheduled.

Availability Constraints

The availability of each clinic-based worker, $s \in \mathcal{S}$, needs to be considered; this includes approved annual leave, scheduled days of training, and working day patterns.

$$\sum_{m \in \mathcal{M}} \sum_{i \in \mathcal{I}^{\text{standard}}} y_{s,m,i,t} \leq a_{s,t} \qquad \forall s \in \mathcal{S},\ t \in \mathcal{T} \tag{4.15}$$

The above constraint ensures that any worker is only assigned to work any given clinic $i \in \mathcal{I}^{\text{standard}}$ on day $t \in \mathcal{T}$ if they are available to work on day t.

Feasibility Constraints

We include the following constraints to ensure that the output schedule is feasible.

Constraint 4.16 ensures that any worker $s \in \mathcal{S}$ is only assigned to a clinic $i \in \mathcal{I}$ on day $t \in \mathcal{T}$ if clinic i is open on day t.

$$y_{s,m,i,t} \leq x_{i,t} \qquad \forall s \in \mathcal{S},\ m \in \mathcal{M}_s,\ i \in \mathcal{I},\ t \in \mathcal{T} \tag{4.16}$$

Constraint 4.17 ensures that each worker $s \in \mathcal{S}$ is only assigned to at most one clinic $i \in \mathcal{I}$ per day $t \in \mathcal{T}$.

$$\sum_{m \in \mathcal{M}_s} \sum_{i \in \mathcal{I}} y_{s,m,i,t} \leq 1 \qquad \forall t \in \mathcal{T},\ s \in \mathcal{S} \tag{4.17}$$

Region Constraints

Each worker has a base region, usually the same region that they live in, and should only be assigned to clinics that are within this same region to ensure that there are reasonable travel times to and from work. Let $r' \in \mathcal{R}$ denote a region different to region $r \in \mathcal{R}$.

$$\sum_{m \in \mathcal{M}_s} \sum_{t \in \mathcal{T}} y_{s,m,i_r,t} = 0 \qquad \forall s \in \mathcal{S}_{r'},\ i \in \mathcal{I}^{\text{standard}},\ r \in \mathcal{R} \setminus r' \tag{4.18}$$

Staff Preference Constraints

Each worker has the option to work overtime or not. The following constraint ensures that workers who do not work overtime are not scheduled any hours in excess of their contracted amount.

$$p_s^{\text{overtime}} \cdot \sum_{m \in \mathcal{M}_s} \sum_{i \in \mathcal{I}} \sum_{t \in \mathcal{T}} y_{s,m,i,t} \leq h_s \cdot |\mathcal{W}| \qquad \forall s \in \mathcal{S} \tag{4.19}$$

Each worker should only be assigned to clinics that align with any stated preferences regarding clinic type. Constraints (4.20) ensures that only workers that are willing to work a full clinic tour are scheduled to do so.

$$p_s^{\text{tour}} \cdot \sum_{m \in \mathcal{M}_s} \sum_{i^{\text{tour}} \in \mathcal{I}} \sum_{t \in \mathcal{T}} y_{s,m,i_{\text{tour}},t} = 0 \qquad \forall s \in \mathcal{S} \tag{4.20}$$

Working Regulation Constraints

In the U.K., working an excess of an average of 48 hours per week is prohibited by law. Therefore, we include the following constraints to ensure that all scheduled work is legal.

$$\sum_{t \in \mathcal{T}} \sum_{i \in \mathcal{I}} \sum_{m \in \mathcal{M}_s} y_{s,m,i,t} \cdot l_{i,t} \leq 48 \cdot |\mathcal{W}| \qquad \forall s \in \mathcal{S} \tag{4.21}$$

The WBS imposes a rule that any clinic-based worker that works 12 hours or longer in one day is required to have the following day off. Thus we utilise constraints (4.22) to ensure this holds. Similarly, the WBS also imposes a requirement that any clinic-based worker due to work a clinic tour must have a day off on the day prior to the tour commencing, and constraints (4.23) ensure this occurs.

$$\sum_{m \in \mathcal{M}_s} y_{s,m,\tilde{i},t} + \sum_{i \in \mathcal{I}^{\text{standard}} \setminus \tilde{i}} \sum_{m \in \mathcal{M}_s} y_{s,m,i,t+1} \leq 1 \qquad \forall s \in \mathcal{S},\ t \in \mathcal{T}\ \tilde{i} \in \mathcal{L}_{12} \tag{4.22}$$

$$\sum_{m \in \mathcal{M}_s} y_{s,m,i_1,t} + \sum_{i \in \mathcal{I}^{\text{standard}}} \sum_{m \in \mathcal{M}_s} y_{s,m,i,t-1} + \sum_{m \in \mathcal{M}_s} y_{s,m,i^{\text{training}},t-1} \leq 1$$
$$\forall t \in \mathcal{T} \setminus t_1,\ s \in \mathcal{S},\ i_1 \in \mathcal{I}^{\text{tours}} \tag{4.23}$$

Tour Constraints

Tour clinics require the same workers to work through the duration of the tour i.e. the same workers need to be scheduled to each clinic of a given tour. Let the number of days of a given tour be denoted by $K \in \mathbb{N}$, such that each clinic day within a tour is denoted by $k \in \mathbb{N}_{\leq K}$. We ensure that if worker $s \in \mathcal{S}$ is scheduled to work a clinic $i_k \in \mathcal{I}^{\text{tours}}$ within a tour, then they are scheduled to work every other clinic within the same tour.

$$\sum_{m\in\mathcal{M}_s} y_{s,m,i_1,t} = \sum_{m\in\mathcal{M}_s} y_{s,m,i_k,t+k} \qquad \forall\, k\in\mathbb{N}_{\le K},\; t\in\mathcal{T},\; i_1\in\mathcal{I}^{\text{tours}},\; s\in\mathcal{S} \tag{4.24}$$

Dummy Clinic Constraints

In order to include worker hours for training and annual leave for the workforce, we utilise dummy clinics. Each worker $s \in \mathcal{S}$ has an associated annual leave dummy clinic, $i_s^{\text{leave}} \in \mathcal{I}$, with length of day, $l_{i,1}$, equivalent to one day of annual leave based on the contracted hours of worker s. Similarly, there is one dummy training clinic, i^{training}, as training days are the same length for all workers.

$$\sum_{m\in\mathcal{M}_s} y_{s,m,i_s^{\text{leave}},t} = g_{s,t}^{\text{leave}} \qquad \forall s\in\mathcal{S},\; t\in\mathcal{T} \tag{4.25}$$

Constraints 4.25 ensures that each worker $s \in \mathcal{S}$ is assigned to their corresponding dummy clinic $i_s^{\text{leave}} \in \mathcal{I}$ for each day of booked annual leave. Likewise, constraints 4.26 ensures that each worker $s \in \mathcal{S}$ is assigned to the dummy training clinic, $i^{training} \in \mathcal{I}$, for each day of booked training.

$$\sum_{m\in\mathcal{M}_s} y_{s,m,i^{\text{training}},t} = g_{s,t}^{\text{training}} \qquad \forall s\in\mathcal{S},\; t\in\mathcal{T} \tag{4.26}$$

Workload Balancing Constraints

We introduce a balancing constraint, involving decision variables $z_{s,w}^{\text{over}}$ and $z_{s,w}^{\text{under}}$, to enable these variables to have the desired values of weekly overtime and 'undertime' respectively, for each worker $s \in \mathcal{S}$.

$$\left(\sum_{t\in\mathcal{T}_w} y_{s,m,i,t}\cdot l_{i,t}\right) - h_s - z_{s,w}^{\text{over}} + z_{s,w}^{\text{under}} = 0 \qquad \forall w\in\mathcal{W},\; s\in\mathcal{S} \tag{4.27}$$

We also utilise the following constraints to determine the bounds of decision variables, $z_{s,w}^{\text{over}}$ and $z_{s,w}^{\text{under}}$. Constraints (4.28) ensure that weekly overtime is non-negative. Constraints (4.29) ensure that weekly 'undertime' for a worker is at most their weekly contracted hours, h_s, and non-negative.

$$0 \leq z^{\text{over}}_{s,w} \tag{4.28}$$

$$0 \leq z^{\text{under}}_{s,w} \leq h_s \tag{4.29}$$

Linking Constraints

To ensure that decision variables $v_{1,s}$ and $v_{2,s}$ are correctly utilised as overtime pay indicators, we have the following constraints.

$$\left(\sum_{m \in \mathcal{M}_s} \sum_{i \in \mathcal{I}} \sum_{t \in \mathcal{T}} y_{s,m,i,t} \cdot l_{i,t} \right) - v_{1,s} \leq C \cdot |\mathcal{W}| \qquad \forall s \in \mathcal{S} \tag{4.30}$$

$$\left(\sum_{m \in \mathcal{M}_s} \sum_{i \in \mathcal{I}} \sum_{t \in \mathcal{T}} y_{s,m,i,t} \cdot l_{i,t} \right) - v_{2,s} \leq h_s \qquad \forall s \in \mathcal{S} \tag{4.31}$$

$$v_{1,s} \in \mathbb{R}_{\geq 0} \qquad \forall s \in \mathcal{S} \tag{4.32}$$

$$v_{2,s} \in \mathbb{R}_{\geq 0} \qquad \forall s \in \mathcal{S} \tag{4.33}$$

4.4 Model Modifications

The original Blood Collection Workforce Scheduling Model provided poor results (as displayed in Appendix C), with scheduled overtime and an unbalanced assignment of clinic hours across all clinic-based workers. With some workers assigned almost all of their contracted hours, others are assigned very few working hours. This is unrealistic and unsustainable as a staffing model, and does not align with the WBS goal of utilising all contracted hours. To improve the fairness of allocated working hours across the clinic workforce, we make the following modifications to the model.

4.4.1 Parameters

In addition to all parameters previously introduced in Table 4.5, we introduce the following:

- **Maximum available clinic hours:** Let $\Omega_{s,w}$ denote the maximum available clinic hours that worker $s \in \mathcal{S}$ could be assigned to in week $w \in \mathcal{W}$. This considers all scheduled clinics in the corresponding region for worker s, and deducts all unavailable clinic hours for worker s due to pre-determined annual leave and training, or their established working day pattern.

- **Total Leave Hours:** Let $c_{s,w}$ denote the total booked annual leave and training hours for worker $s \in \mathcal{S}$ on week $w \in \mathcal{W}$. Therefore, let $\mathcal{S}^{c_{s,w}<0.5\cdot h_s} \subset \mathcal{S}$ denote the set of clinic workers that have fewer than 50% of their contracted hours, h_s, scheduled as annual leave and/or training.

With all new parameters introduced, we continue to the modified objective functions of the improved model.

4.4.2 Objective Function

To enable the model to consider a fairness of assignment of hours across workers, we utilise the decision variables $z_{s,w}^{\text{over}}$ and $z_{s,w}^{\text{under}}$ i.e. weekly overtime and 'undertime' per worker respectively. Overtime is calculated by the WBS over the whole four-week planning horizon. However, since these variables consider each worker's assigned hours per week, when considering across all weeks in the planning horizon, this gives an indication of how far the assigned hours are from the target of a worker's contracted hours and how well balanced they are over the planning horizon.

Therefore, we formulate the two modified objective functions as follows:

Objective Function One

$$\text{Minimise} \sum_{s \in \mathcal{S}} \left(b_s \cdot v_{2,s} \sum_{m \in \mathcal{M}_s} \sum_{i \in \mathcal{I}} \sum_{t \in \mathcal{T}} y_{s,m,i,t} \cdot l_{i,t} + \frac{b_s}{2} \cdot v_{1,s} \sum_{m \in \mathcal{M}_s} \sum_{i \in \mathcal{I}} \sum_{t \in \mathcal{T}} y_{s,m,i,t} \cdot l_{i,t} \right)$$
$$+ \beta_1 \sum_{s \in \mathcal{S}} \sum_{w \in \mathcal{W}} z^{\text{over}}_{s,w} + \beta_2 \sum_{s \in \mathcal{S}} \sum_{w \in \mathcal{W}} z^{\text{under}}_{s,w} + \beta_3 \sum_{s \in \mathcal{S}_{\text{dep}}} \left(\sum_{m \in \mathcal{M}_{s,\text{sup}}} \sum_{i \in \mathcal{I}} \sum_{t \in \mathcal{T}} y_{s,m,i,t} \right) \quad (4.34)$$

Objective Function Two

$$\text{Minimise} \sum_{s \in \mathcal{S}} \left(b_s \sum_{m \in \mathcal{M}_s} \sum_{i \in \mathcal{I}} \sum_{t \in \mathcal{T}} y_{s,m,i,t} \cdot l_{i,t} + \frac{b_s}{2} \cdot v_{1,s} \sum_{m \in \mathcal{M}_s} \sum_{i \in \mathcal{I}} \sum_{t \in \mathcal{T}} y_{s,m,i,t} \cdot l_{i,t} \right)$$
$$+ \beta_1 \sum_{s \in \mathcal{S}} \sum_{w \in \mathcal{W}} z^{\text{over}}_{s,w} + \beta_2 \sum_{s \in \mathcal{S}} \sum_{w \in \mathcal{W}} z^{\text{under}}_{s,w} + \beta_3 \sum_{s \in \mathcal{S}_{\text{dep}}} \left(\sum_{m \in \mathcal{M}_{s,\text{sup}}} \sum_{i \in \mathcal{I}} \sum_{t \in \mathcal{T}} y_{s,m,i,t} \right) \quad (4.35)$$

We use β_1, β_2, and $\beta_3 \in \mathbb{R}$ as constants to allow the smaller terms of the objective function to be comparable to the first larger term.

By including the minimisation of the total weekly overtime and total weekly 'undertime' across all clinic-based workers, the model is motivated to assign the target weekly hours to each worker i.e. their contracted hours.

4.4.3 Constraints

To ensure that no worker is scheduled an unreasonable amount of overtime, we include the following constraints to limit weekly overtime to 50% of a worker's contracted hours.

$$z^{\text{over}}_{s,w} \leq 0.5 \cdot h_s \qquad \forall s \in \mathcal{S},\ w \in \mathcal{W} \quad (4.36)$$

Additionally, we include new constraints to reduce the likelihood of a worker being scheduled hours below a lower threshold.

$$\sum_{t\in\mathcal{T}_w}\sum_{i\in\mathcal{I}^{\text{standard}}}\sum_{m\in\mathcal{M}_s} y_{s,m,i,t}\cdot l_{i,t} \geq 0.5\cdot\min\{h_s - c_{s,w}, \Omega_{s,w}\} \qquad \forall s\in\mathcal{S}^{c_{s,w}<0.5\cdot h_s},\ w\in\mathcal{W} \tag{4.37}$$

Constraints (4.37) ensure that for each worker with at least half of their contracted hours available for clinic hours on a given week $w \in \mathcal{W}$, the minimum of either half of their remaining available contracted hours $(h_s - c_{s,w})$ or half of the maximum available clinic hours that week (Ω), are assigned to worker s. We include the latter value, $\Omega_{s,w}$ to ensure that the problem remains feasible if there are no clinics scheduled in a given week.

4.5 Summary

In this chapter, the current practice of clinic-based workforce scheduling at the Welsh Blood Service has been introduced, including details about how the workforce are organised, and limitations of this current practice have been identified. A formulation of the Blood Collection Workforce Scheduling Problem (BCWSM) is presented in the form of a linear programme to assign workers to clinics optimally. We detailed all parameters of the problem, followed by two alternative objective functions; minimisation of overtime costs, and minimisation of the cost of all scheduled hours. Both objective functions also include a penalty function to discourage the model from selecting a Deputy Supervisor to be assigned to the clinic role of a Supervisor if there are Supervisors available. We introduced all of the constraints required to ensure the output of the model is a feasible solution for the WBS. Additionally, we identified areas of improvement for the formulation and introduced two modified alternative objective functions to include the minimisation of both weekly overtime and 'undertime' per worker, and introduced a new constraint. The aim of this modified version of the model is to create a more realistic schedule output for the Welsh Blood Service.

In the next chapter, Chapter 5, we discuss the development of a prototype decision support tool for the BDCSP, alongside the development of both the model described in this chapter (BCWSM), and the clinic scheduling model presented in Chapter 3 (BDCSM).

Chapter 5

Model Development

This chapter discusses the development of both the Blood Donation Clinic Scheduling Model (BDCSM) and the Blood Collection Workforce Scheduling Model (BCWSM), in addition to a prototype decision support tool for the Blood Donation Clinic Scheduling Problem (BDCSP). This chapter is structured as follows: the background of the models and their development is briefly discussed in Section 5.1, with a description of the available data in the subsequent Section 5.2. COIN-OR and its CBC solver are described in Section 5.3, with Python and PuLP outlined afterwards in Section 5.4. Section 5.5 presents the prototype tool created in Microsoft Excel to solve the BDCSM, while sections 5.6 and 5.7 detail the development of the models written in Python for the BDCSP and BCWSP, respectively.

5.1 Background

With the initial goal of this research project to create a decision support tool that would be implemented at the Welsh Blood Service (WBS), a prototype clinic scheduling tool was developed in Microsoft Excel, with the aim to trial this tool on-site at the WBS Headquarters with the planning team. Unfortunately, due to the COVID-19 pandemic, this tool was not able to be piloted, and to adapt to the new circumstances, the focus became furthering the development and improvement of the BDCSM and BCWSM in Python (utilising the Python package PuLP as a solver) to ensure that these models are feasible for the WBS and a realistic representation of the clinic scheduling process and limitations.

5.2 Available Data

The Welsh Blood Service provided all available data required for the clinic scheduling process. This includes all 'bleed figures' i.e. blood collection data, since January 2017; these figures include how many units of blood were collected from each clinics, how many of these were viable donations, and how many donors attended. Similarly, all blood issuing data was provided by the WBS for the period from January 2017 to March 2020, which includes the age of each blood product at the time of issue to a hospital. Due to the COVID-19 pandemic and its effect on the operations of the WBS donation clinics, collection data from 2020 is not a true representation of business as usual and therefore we consider only data prior to this.

All clinic data was made available, such as each clinic's location, donor capacity, staffing level requirements, and typical availability pattern. Additionally, the WBS provided an anonymised dataset of the donation clinic workforce, with identifiable information such as names and addresses removed. This data includes staff roles, skill level, agreed working day patterns and contracted hours. While only salary bands for each employee were given by the WBS and not a specific pay rate, information regarding salary band pay rates and annual leave entitlement is openly accessible online, published by the NHS.

5.3 COIN-OR Solver

All three models utilise COIN-OR CBC optimisation engine, including the prototype donation clinic scheduling tool via the OpenSolver Excel add-in. There are existing scheduling tools that utilise COIN-OR via OpenSolver such as the 'Shift Scheduling Game' developed by Michigan Center for Healthcare Engineering and Patient Safety [37]. COIN-OR solvers are free and open source which is a necessity for the Welsh Blood Service as the minimisation of cost is a priority as an NHS organisation. Although more advanced commercial solvers such as Gurobi and CPLEX are known to have the best performance, these solvers are not an option due to the significant associated costs. However, COIN-OR CLP and CBC solvers are among the strongest open-source solvers for linear programming with the CLP solver concluded to be the top performing open-source solver by the U.S. Office of Scientific and Technical Information [38] when compared with popular solvers GNU Linear Programming Kit (GLPK), lp_solve, and Modular In-core Nonlinear Optimisation System (MINOS).

COIN-OR (Computational Infrastructure for Operations Research) is a project managed by the COIN-OR Foundation with the aim to build an "open-source community for operations research software" [1]. There are many smaller projects within the COIN-OR project, namely the development of various software for a range of problems, methods and coding languages. There are two main linear programming solvers developed by the COIN-OR community; CLP (COIN-OR Linear Programming) which mainly uses the simplex algorithm as its main algorithm, and CBC (COIN-OR Branch and Cut) which is a mixed integer linear programme-based branch and cut library and also utilises CLP. Both of these solvers are written in C++ but can be utilised with various languages through available packages such as PuLP. Due to the nature of the BDCSP with decision variables required to have integer values, the CBC solver is the more suitable choice.

5.4 Python and PuLP

PuLP is a linear programme modeller written in the computer programming language Python [2]. It can utilise a number of solvers such as CPLEX, GUROBI and COIN-OR CLP/CBC. Due to the minimisation of cost being important to the Welsh Blood Service, our developed models utilise PuLP as it is open source and therefore free to use, and is simply installed as a Python package. We choose Python as the primary language for development of the model as it is open-source, readable, high level and thus easy to learn [80]. This supports the possibility for future implementation of the model (embedded within a tool) at the WBS and indeed for maintenance, expansion and further development of the model by the WBS themselves.

5.5 Development of a Blood Donation Clinic Scheduling Decision Support Tool

With the initial goal outcome of this research being an implementable tool for the Welsh Blood Service to optimise their clinic and workforce scheduling, we developed a prototype clinic scheduling decision support tool. In order to make ongoing maintenance of the tool as straightforward as possible, in addition to maximising the usability of the tool among the clinic planning staff, the tool was created in Microsoft Excel. The tool utilises OpenSolver to solve the Blood Donation Clinic Scheduling Problem (BDCSP) optimally using COIN-OR CBC optimization engine. It was planned to pilot the tool at the WBS with an iterative feedback approach from clinic planning staff, and to use this feedback to further improve and adjust the tool as required. However, due to the COVID-19 pandemic, the tool was not able to be piloted, and thus a second stage of the prototype tool for optimisation of workforce scheduling was not developed.

5.5.1 Aims of Decision Support Tool

The development of a tool for the Welsh Blood Service to assist their clinic planning process would not only minimise clinic associated costs and improve efficiency of the collection of blood, but also improve efficiency within the planning department.

Currently, the clinic planning process is extremely tedious and time consuming, with an initial four-week clinic schedule taking an average of five working days to create. The use of such a tool would produce an optimal clinic schedule in significantly less time and also with less input required from the planning team.

Additionally, the tool needs to be easily accessible, inexpensive, and usable for the planning team with minimal training required. The maintenance of the tool is also required to be minimal, as the IT department of the WBS have limited resources. This led to the decision for the tool to be embedded within Microsoft Excel, as it is a software that all staff at the WBS are widely familiar with and would require minimal training to use the tool.

5.5.2 Decision Support Tool Development

The prototype considers all blood donation clinics in the South West region (i.e. those associated to the Dafen base) which totals 51 clinics. This region was chosen as the significantly larger South East region poses a more complex problem with up to four clinics permitted to be scheduled each day, whereas the three remaining regions all have resource levels to support a maximum of one clinic per day. The South West region presents a slightly more complex problem than both North East and North West regions as multiday clinics occur in this region. This provides the opportunity to incorporate constraints related to these clinics that would not be relevant in either of the north regions.

The decision support tool consists of a Microsoft Excel file with multiple sheets for details such as availability and estimated supply. Each type of parameter has its own individual sheet such as a sheet for each type of availability, and named clearly to ensure the clinic planner is able to easily access data as required. Due to the limitations of the software, all data to be considered in the optimisation model must be on the same sheet, and thus to prevent important data accidentally being altered during the pilot study, all data is stored in alternative protected sheets, and simply called upon on the 'model' sheet. Since the model is developed to be specifically for the WBS, all required clinic data is already input and therefore, most data is unlikely to require editing except for cases where a new clinic is added or the details of an existing clinic have changed. Therefore, the 'model' sheet is located first, with the aim that this sheet, along with the adjacent 'schedule' sheet, are the only sheets that require frequent use. Of the remaining sheets, those that are more likely to require altering such as 'clinic frequency' are located before sheets that require very infrequent adjustments such as 'clinic types' which categorises clinics into one-day clinics and multi-day clinics.

Planning Horizon

The WBS plan their clinics over four-week periods at a time; to ensure the tool is as implementable as possible, a four-week planning horizon is considered with the 'start date' of the planning horizon easily adjusted on the 'model' sheet. The specific planning horizon considered for the pilot version of the tool is 06/01/2020 - 02/02/2020.

Clinic Availability

Due to the nature of mobile clinics, availability of clinic venues for the WBS is a major constraint. In most cases, clinic availability is predetermined and follows a regular pattern. These availability patterns can be divided into three types; week-day availability (e.g. Mondays only), seasonal availability (e.g. no winter months, or school holidays only) and frequency-based availability (ensuring that a pre-determined minimum number of days have lapsed between consecutive occurrences of a given clinic).

To inform the optimisation model of the availability of each clinic, a binary clinic availability matrix is considered for each type of availability, where a '1' represents that the corresponding clinic is available on the corresponding day of the planning horizon, and '0' represents that the clinic is unavailable. These matrices take the form of a table on an individual sheet per availability type, with clinic names in the leftmost column, availability patterns in the adjacent column, followed by the binary matrix with each day of the planning horizon as the column headers, as displayed in the example for weekday availability in Figure 5.1.

	A	B	C	D	E	F	G	H	I
1		Day of Planning Horizon	1	2	3	4	5	6	7
2		Day of Week	1	2	3	4	5	6	7
3		Day of Month	6	7	8	9	10	11	12
4	**Clinics**	**Availability Pattern**	**06/01/2020**	**07/01/2020**	**08/01/2020**	**09/01/2020**	**10/01/2020**	**11/01/2020**	**12/01/2020**
5	Aberaeron	Weekdays Only	1	1	1	1	1	0	0
6	Ammanford		1	1	1	1	1	0	0
7	Burry Port		0	1	0	0	1	0	0
8	Cardigan		1	0	0	1	0	0	0
9	Carmarthen		1	1	1	1	1	0	0
10	Clydach		1	0	1	0	1	0	0
11	Crymych	Weekdays Only	1	1	1	1	1	0	0
12	Fishguard	Weekdays Only	1	1	1	1	1	0	0
13	Gendros	Wednesday Only	0	0	1	0	0	0	0
14	Glynneath	Weekdays Only	1	1	1	1	1	0	0
15	Gorseinon	Friday Only	0	0	0	0	1	0	0
16	Gwaun Cae Gurwen	Friday Only	0	0	0	0	1	0	0
17	Haverfordwest	Wednesday, Thursday Only	0	0	1	1	0	0	0
18	Kidwelly Community Hall	Friday Only	0	0	0	0	1	0	0
19	Killay	Weekdays Only	1	1	1	1	1	0	0
20	Lampeter	Monday, Tuesday, Wednesday, Thursday On	1	1	1	1	0	0	0
21	Liberty Stadium	Monday Only	1	0	0	0	0	0	0
22	Llandeilo	Weekdays Only	1	1	1	1	1	0	0
23	Llandovery	Friday Only	0	0	0	0	1	0	0
24	Llandysul	Weekdays Only	1	1	1	1	1	0	0
25	Llanelli	Monday Only	1	0	0	0	0	0	0
26	Milford Haven	Tuesday, Wednesday, Thursday, Friday Only	0	1	1	1	1	0	0
27	Mumbles	Weekdays Only	1	1	1	1	1	0	0
28	Murton	Friday Only	0	0	0	0	1	0	0
29	Narberth	Tuesday, Wednesday, Thursday, Friday Only	0	1	1	1	1	0	0
30	Newcastle Emlyn	Monday, Tuesday, Wednesday Only	1	1	1	0	0	0	0
31	New Quay	Monday, Tuesday, Wednesday, Thursday On	1	1	1	1	0	0	0
32	Neyland	Weekdays Only	1	1	1	1	1	0	0
33	Pembroke Dock	Weekdays Only	1	1	1	1	1	0	0
34	Pembroke	Weekdays Only	1	1	1	1	1	0	0
35	Penclawdd	Friday Only	0	0	0	0	1	0	0
36	Penygroes (Llanelli)	Weekdays Only	1	1	1	1	1	0	0
37	Pontardawe	Tuesday Only	0	1	0	0	0	0	0
38	Pontyberem	Thursday Only	0	0	0	1	0	0	0

Weekdays Only
Monday Only
Tuesday Only
Wednesday Only
Thursday Only
Friday Only
Monday, Tuesday Only
Monday, Wednesday Only

model | schedule | clinic_frequency | seasonal_availability | **weekday_availability** | availability_patterns | constraints | expected_supply

Figure 5.1: Prototype Tool - Weekday Availability of Clinics

Visual Basic (VBA) code has been utilised to provide a drop-down menu to select the availability pattern of a clinic, as shown in Figure 5.1, for both weekday and seasonal availability. Once an option is selected, the row automatically generates the associated availability (in the form of a binary value per day of the planning horizon) for that clinic. This matrix is then referred to by the model to ensure each clinic is only scheduled on days where the venue is available.

Frequency-based availability is determined by observing the most recent date at which each clinic was held, prior to the current planning horizon. These dates are input into the table for frequency-based availability, with the adjacent column containing the date at which each clinic is next available to be scheduled, by considering the corresponding maximum frequency for a given clinic. Each cell in the binary matrix for the availability then utilises a formula that considers this next available date alongside the given date of the planning horizon (denoted by the column headers), and inputs a '1' if a given clinic is available to be scheduled on a given day, and '0' otherwise.

Multiday Clinics

The Welsh Blood Service operates multiday clinics in several locations across Wales. These are usually in locations where the donor panel is large and the venue has limited availability throughout the year such as a university. The duration of these clinics vary from two to five consecutive days, with some multiday clinics in the South East region occurring over non-consecutive days.

	A	B	C	D	E	F	G	H	I
72	**Consecutive Day Clinics Start and Stop Days**		1	2	3	4	5	6	7
73	**Start/Stop**	**Clinic**	**06/01/2020**	**07/01/2020**	**08/01/2020**	**09/01/2020**	**10/01/2020**	**11/01/2020**	**12/01/2020**
74	Start	Aberystwyth University	0	0	0	0	0	0	0
75	Stop	Aberystwyth University	0	0	0	0	0	0	0
76	Start	Aberystwyth	0	0	0	0	0	0	0
77	Stop	Aberystwyth	0	0	0	0	0	0	0
78	Start	City and County of Swansea Council	0	1	0	0	0	0	0
79	Stop	City and County of Swansea Council	0	0	1	0	0	0	0
80									
81	**Consecutive Clinics Constraints**	**Required Duration**	**Scheduled Duration**	**Start Day**	**Stop Day**	**Is clinic scheduled?**	**Sum of Start & Stop Days**	**Constraint**	
82	Aberystwyth University	3	1	0	0	0	0	0	
83	Aberystwyth	3	1	0	0	0	0	0	
84	City and County of Swansea Council	2	2	2	3	1	2	2	

Figure 5.2: Prototype Tool - Multiday Clinics

Configuring these clinics in the Excel tool requires a different method to the standard one-day clinics. We created 'start' and 'stop' days (both binary decision variables) for each multiday clinic, with the former being the first day that the clinic is to be operated, and the latter being the last day of the clinic. We also created a dummy (binary) decision variable per multiday clinic for if the clinic is to be scheduled during the planning horizon or not. With the inclusion of constraints ensuring that the number of days between (and including) the start and stop days is equal to the required duration of the multiday clinic, the model successfully includes clinics of this type within the schedule.

Supply and Demand

The supply estimates for each clinic are calculated based on the mean of the three previous 'bleed' figures for the clinic, as this is the method that the WBS used prior to the COVID-19 pandemic. The tool allows for the weekly demand to be input, and ensures that the sum of these per week is met for each week of the planning horizon.

The weekly demand input into the tool is set at an arbitrary value of 372 units of blood; this is based on 21.3% of the total weekly demand across all regions, as this percentage is the mean proportion of total blood donations collected from the South West region during 2019. The total weekly demand across all regions considered for the tool is the value of 1750 units, which is the value that the clinic planning team frequently use as a target, based on rough estimates by the WBS. A constraint is created for each week's estimated supply to ensure that the weekly demand is met.

Decision Variables and Objective Function

The decision variables considered in this prototype tool are as described in Section 3.3, with $x_{i,t}$ as binary decision variables where $i \in \mathcal{I}$ denotes all clinics in the South West region and $t \in \mathcal{T}$ denotes the days of the given planning horizon. For the multiday clinics, each start and stop day are individual binary decision variables.

The tool considers only one objective function: the minimisation of overcollection i.e. the minimisation of blood collected in excess to demand. This is due to the tool being developed prior to the proposition to include an alternative objective function to minimise the total number of scheduled clinic days over a given planning horizon to promote a reduction of monetary costs.

Excluded Constraints and Simplifications

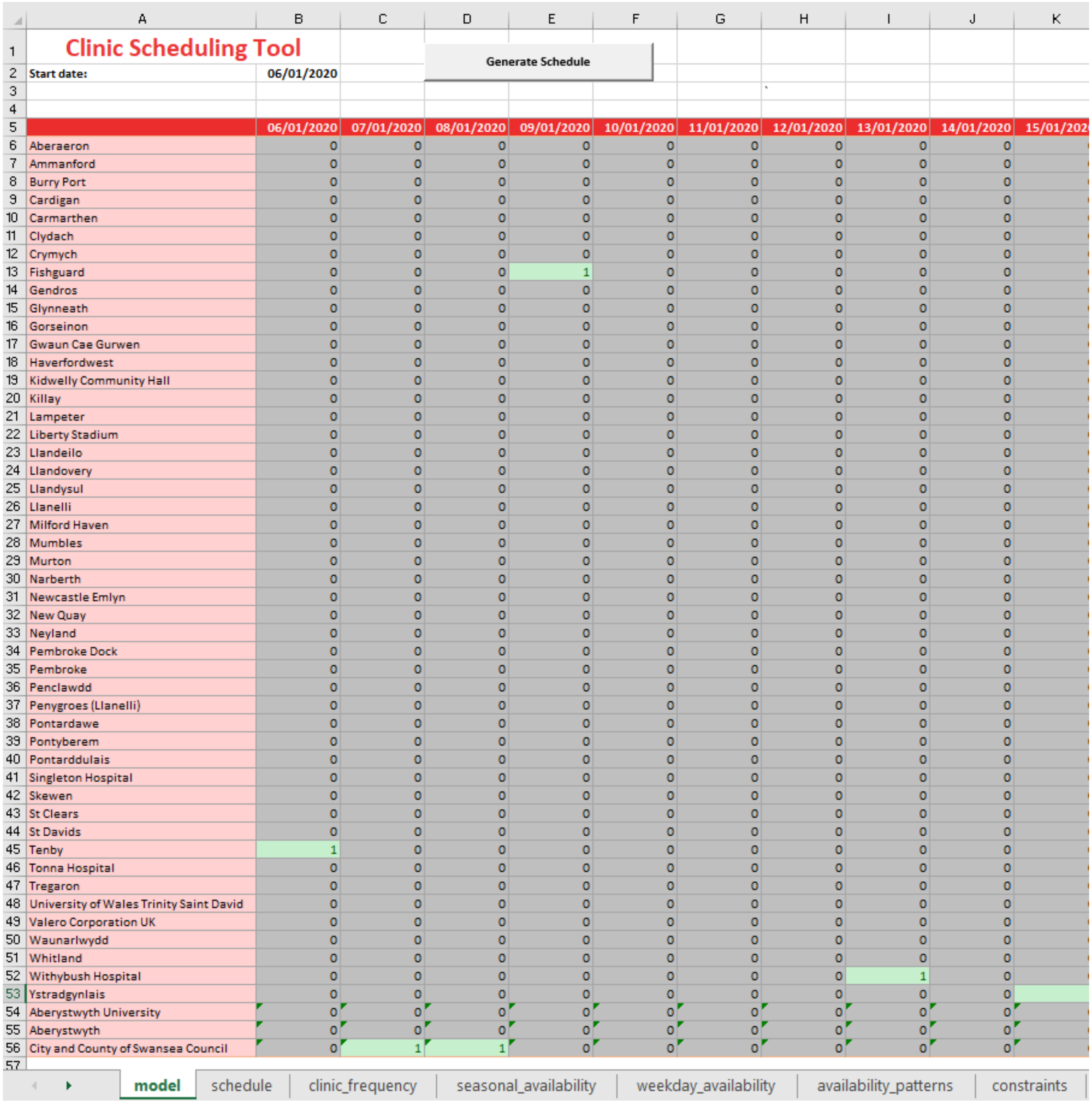
With the initial purpose of the prototype tool to be trialled at the WBS and further developed as an iterative process, not every constraint detailed in Section 3.3 is included. Constraints (4.22) and (3.26) are not included in the prototype tool as a simplification of the problem for an initial pilot of the tool; these constraints ensure that the number of clinics scheduled in a region are restricted if a day off for some workers is required, as a result of either a 12-hour working day or a clinic tour. For the same reasons, constraints (3.18) - (3.20) and (3.21) are not included in the tool.

Similarly, constraints (3.23) and (3.24) are excluded for simplification; although there are technically clinic tours in the region, they consist of several consecutive days at the same clinic venue and thus the clinic duration enforces all days of the tour to be scheduled. However, constraints (3.27) are not applicable to the South West region as there are no obligatory clinics in the region.

Tool Functionality and Output

The prototype considers all blood donation clinics in the South West region (i.e. those associated to the Dafen base) and solves within one minute. The clinic scheduling model can be solved by the user by inputting the first day of the four-week planning horizon and selecting the embedded 'Generate Schedule' button (see Figure 5.3). The scheduled clinics display in green in the respective date column for each day that the clinic is to operate. Alternatively, each scheduled clinic is automatically listed under the corresponding date on the 'schedule' sheet of the file.

The focus with this tool was to ensure that it is user-friendly and accessible. Many options provide automated outputs for ease of use, such as choosing the start date of the planning horizon and availability patterns of each clinic. In addition to the scheduled clinics highlighted in green, the sheet named 'schedule' contains all scheduled clinics listed in columns titled with the corresponding date of a scheduled clinic. This schedule sheet is automatically generated using a macro written in VBA to provide the clinic planner with a clear clinic schedule. Figure 5.4 presents an example of this output, displaying the first week of the planning horizon.



	A	B	C	D	E	F	G	H	I	J	K
1	Clinic Scheduling Tool			Generate Schedule							
2	Start date:	06/01/2020									
3											
4											
5		06/01/2020	07/01/2020	08/01/2020	09/01/2020	10/01/2020	11/01/2020	12/01/2020	13/01/2020	14/01/2020	15/01/202
6	Aberaeron	0	0	0	0	0	0	0	0	0	
7	Ammanford	0	0	0	0	0	0	0	0	0	
8	Burry Port	0	0	0	0	0	0	0	0	0	
9	Cardigan	0	0	0	0	0	0	0	0	0	
10	Carmarthen	0	0	0	0	0	0	0	0	0	
11	Clydach	0	0	0	0	0	0	0	0	0	
12	Crymych	0	0	0	0	0	0	0	0	0	
13	Fishguard	0	0	0	1	0	0	0	0	0	
14	Gendros	0	0	0	0	0	0	0	0	0	
15	Glynneath	0	0	0	0	0	0	0	0	0	
16	Gorseinon	0	0	0	0	0	0	0	0	0	
17	Gwaun Cae Gurwen	0	0	0	0	0	0	0	0	0	
18	Haverfordwest	0	0	0	0	0	0	0	0	0	
19	Kidwelly Community Hall	0	0	0	0	0	0	0	0	0	
20	Killay	0	0	0	0	0	0	0	0	0	
21	Lampeter	0	0	0	0	0	0	0	0	0	
22	Liberty Stadium	0	0	0	0	0	0	0	0	0	
23	Llandeilo	0	0	0	0	0	0	0	0	0	
24	Llandovery	0	0	0	0	0	0	0	0	0	
25	Llandysul	0	0	0	0	0	0	0	0	0	
26	Llanelli	0	0	0	0	0	0	0	0	0	
27	Milford Haven	0	0	0	0	0	0	0	0	0	
28	Mumbles	0	0	0	0	0	0	0	0	0	
29	Murton	0	0	0	0	0	0	0	0	0	
30	Narberth	0	0	0	0	0	0	0	0	0	
31	Newcastle Emlyn	0	0	0	0	0	0	0	0	0	
32	New Quay	0	0	0	0	0	0	0	0	0	
33	Neyland	0	0	0	0	0	0	0	0	0	
34	Pembroke Dock	0	0	0	0	0	0	0	0	0	
35	Pembroke	0	0	0	0	0	0	0	0	0	
36	Penclawdd	0	0	0	0	0	0	0	0	0	
37	Penygroes (Llanelli)	0	0	0	0	0	0	0	0	0	
38	Pontardawe	0	0	0	0	0	0	0	0	0	
39	Pontyberem	0	0	0	0	0	0	0	0	0	
40	Pontarddulais	0	0	0	0	0	0	0	0	0	
41	Singleton Hospital	0	0	0	0	0	0	0	0	0	
42	Skewen	0	0	0	0	0	0	0	0	0	
43	St Clears	0	0	0	0	0	0	0	0	0	
44	St Davids	0	0	0	0	0	0	0	0	0	
45	Tenby	1	0	0	0	0	0	0	0	0	
46	Tonna Hospital	0	0	0	0	0	0	0	0	0	
47	Tregaron	0	0	0	0	0	0	0	0	0	
48	University of Wales Trinity Saint David	0	0	0	0	0	0	0	0	0	
49	Valero Corporation UK	0	0	0	0	0	0	0	0	0	
50	Waunarlwydd	0	0	0	0	0	0	0	0	0	
51	Whitland	0	0	0	0	0	0	0	0	0	
52	Withybush Hospital	0	0	0	0	0	0	0	1	0	
53	Ystradgynlais	0	0	0	0	0	0	0	0	0	
54	Aberystwyth University	0	0	0	0	0	0	0	0	0	
55	Aberystwyth	0	0	0	0	0	0	0	0	0	
56	City and County of Swansea Council	0	1	1	0	0	0	0	0	0	
57											

Figure 5.3: Prototype Tool - Model Solution

Figure 5.4: Prototype Tool - Schedule Output

	A	B	C	D	E	F
1	06 January 2020	07 January 2020	08 January 2020	09 January 2020	10 January 2020	11 January 2020
2	Tenby	City and County of Swansea Council	City and County of Swansea Council	Fishguard		
3						

5.6 Development of the Blood Donation Clinic Scheduling Model in Python

To avoid limitations of VBA and Microsoft Excel, a more detailed and accurate optimisation model for the scheduling of blood donation clinics was created using Python and the linear programme package PuLP. This model follows the full formulation for the BDCSP set out in Chapter 3.

5.6.1 Aims of Model

The BDCSM was built in Python to schedule the WBS clinics optimally, across all regions. The output of the model is a four-week clinic schedule that aims to match supply to demand and/or minimise clinic associated costs by reducing the number of clinics scheduled to meet the demand. This forms an ideal ‘starting point’ for the clinic planning process as an initial schedule, and thus can be changed manually by the planning team if any clinic venues are newly unavailable for the scheduled date.

The benefits of this model being used in place of the current practice at the WBS include a reduction in manual scheduling for the clinic planning team, an ability to increase cost efficiency of blood collection by scheduling as few clinics as possible required to meet demand, and to better match supply to demand. These all contribute to answering to research question one: How can mathematical modelling help to schedule the Welsh Blood Service’s blood donation clinics more efficiently?

The BDCSM forms the first stage of the full scheduling model for the WBS, with the output of the BDCSM to be used as input into stage two of the model (the

BCWSM) to assign donation clinic-based workers to scheduled clinics optimially.

5.6.2 Development of Model

The model was built incrementally, beginning with a small simplified model considering only the South West regions, similar to the prototype tool version. Gradually, the model was expanded to include all constraints detailed in Section 3.3 and to consider all four regions of Wales, to ensure that the model provided an accurate reflection of the clinic scheduling process at the WBS and all related rules and limitations.

Input of Clinic Data

All clinic data that is relevant to the clinic scheduling process was collated into one Microsoft Excel file, separated into individual sheets by region, with an additional sheet with all clinic data included i.e. all regions collectively. For each clinic, this data includes a unique identification key, the number of days that a clinic should run each time it is scheduled, availability patterns, estimated supply, the expected length of working day for each day that a clinic runs, the maximum frequency that a clinic may be scheduled, and the most recent date that a clinic was scheduled prior to specific planning horizons. The model reads this file and utilises a function specifically created to extract relevant data into a new dataset, also known as a 'dataframe' in the Python package, Pandas. This function requires a selected region and season as arguments, and returns the corresponding dataset.

Dummy clinics were created to enable workers to be assigned to a corresponding dummy clinic for any annual leave or training days. A total of 12 dummy clinics were constructed, one for each number of weekly contractual hours for the clinic-based workforce at the WBS. The length of day for these dummy clinics are the corresponding number of contracted hours divided by five (working days) as this is how the WBS calculates the number of hours accounted for by a day of annual leave per worker. One dummy clinic is created for training days as all days of training for their clinic-based workforce are equivalent to 7.5 hours.

Planning Horizon

The planning horizon is input by assigning a date value to the 'initial date' parameter, which is the first day of the planning horizon, and assigning a value to the 'number of weeks' parameter. All dates regarding school and university term dates for the given calendar year are written in the main script of the model, and need to be manually input as date ranges each year to enable them to be considered for clinic availability patterns.

Clinic Availability Patterns

Clinic availability patterns are input using specifically created functions, one for each type of clinic availability pattern i.e. weekday availability and seasonal availability. These functions require a unique clinic key and the planning horizon as arguments, and return a list of binary values to represent the clinic availability i.e. a '0' if the clinic is not available to be scheduled on the given day, or a '1' if the clinic is available. These functions look up the weekday or seasonal availability pattern against the unique clinic key in the clinic dataset, and the specified term dates (if applicable) to determine which days of the planning horizon may be affected.

Clinic Duration Patterns

Clinic duration patterns are written in the main script of the model in the format of $|\mathcal{T}| \times |\mathcal{T}|$ binary matrices using the Python package, NumPy, where $\mathcal{T}$ denotes the set of days in the given planning horizon. Each column of the matrix represents a day of the planning horizon, whilst each row represents which day would be the starting day of a clinic, if selected. An example of how these matrices work is provided in Table 3.2 with examples of the Python code in Appendix D. All of the clinic duration patterns that the WBS used prior to the COVID-19 pandemic are included in the model.

Estimated Supply

The WBS typically calculates their estimated supply per clinic based on an average 'bleed' over the last three times the given clinic occurred i.e. the number of attendees that began a donation. The WBS utilise these figures during clinic scheduling to estimate the supply of a clinic. Since there is usually a small but significant decrease in the number of viable collections from the bleed figures due to issues during the donation that result in a unviable donation, this drop-off rate must be accounted for in supply estimates. Therefore estimated supply is calculated by taking the mean viable collected donations over the three previous clinics (of the same location and type) instead of bled donors to incorporate the typical drop-off rate per clinic.

Estimated Demand

Demand data is manually input in the main script per week per region. Since the WBS are conducting their own research to identify the 'true' demand, the values considered for our experimental results in Chapter 6 are calculated using our own methods, with these calculations to be replaced with the WBS methods once finalised, if the model is to be implemented. To estimate the weekly demand, we used the WBS forecasts of issue of whole blood products per calendar week of 2019 alongside the average age of blood products at date of issue (averaged over all blood types for 2017-2018). By combining these, we worked backwards from week of forecasted issue to estimated week of collection to determine the total number of estimated bags that would be required to be collected each week to meet the forecasted demand with which it corresponds. In addition, an average of 3.9% of viable collections do not make it to inventorised stock due to issues that can arise during production. To account for this, we divide each weekly estimated demand by 96.1%.

Decision Variables, Objective Functions and Constraints

As described in the formulation of the BDCSP in Section 3.3, the model considers both $x_{i,t}$ and Δ_w^+ as decision variables, for a given region's clinics and a given planning horizon. The model considers three alternative objective functions; the minimisation of the total number of clinics days scheduled over the given planning horizon, the minimisation of overcollection, and the minimisation of both the total clinic days scheduled and overcollection. The objective function considered is to be chosen each time the model is run. All constraints detailed in Section 3.3 are included in the model.

Testing and Validation of Results

Throughout the incremental development of the stage one model, the solutions have been inspected to ensure that they are as expected. To ensure that any solution reached by the model meets all of the included constraints, an assert statement is considered for each constraint to ensure that it holds true. This ensures that in circumstances where an optimality gap or time limit is considered during the 'solve' process, the solution is still valid.

We developed a package of functions to enable the model to perform various actions such as cleansing of data, retrieval of weekday and seasonal availability patterns, and retrieving the distance between two given clinics. All of these functions were also developed in Python and were each unit tested to ensure that they act as expected.

5.6.3 Development of Experimental Scenarios

To assess the solutions of the model, and how the model performs under different scenarios, we design various instances for experimental trials of the model.

Test Instances

We designed various instances which can be found in Table 5.1. There were three main factors identified to likely have a significant effect on the goodness of a solution and the computation time: the region(s) considered, the planning horizon, and the inclusion of more specific constraints to reflect the current collection model of the Welsh Blood Service. A more detailed explanation of each of these factors follows.

Regions

Firstly, we consider each region of Wales separately (of which there are four) in addition to the instance considering the whole of Wales. The number of clinics and resource levels vary greatly between regions, with South East having significantly more of both. It was anticipated from early stages of the development of the model that due to the increasing complexity associated with increasing the number of clinics available to be scheduled, the instances considering all regions together would be computationally expensive. It was a possibility that an optimal solution for All Wales instances may not be reached within a feasible time, and that a schedule for the whole nation may need to be comprised of each region schedule when solved individually. Additionally observing the results from these instances of individual regions can provide further insight into the model and indeed the clinic scheduling problem.

Planning Horizon

Secondly, we consider different planning horizons. There are seasonal availability constraints for a significant number of clinics, and so it is important to understand the impact of this variation on the schedule output. We consider two specific four-week planning horizons; one in the winter season (24/01/2019 to 03/02/2019) and one in the summer season (22/07/2019 to 18/08/2019). Both considered planning horizons are four weeks in length and commence on a Monday, as this reflects the current planning process at the Welsh Blood Service.

Table 5.1: Stage One - Base Instances

Instance Name	Region	Number of Clinics in Region	Number of Available Clinics	Planning Horizon Length	Planning Horizon Season	Additional Constraints Included	Mean Time Window per Available Clinic	Total Available Clinic Days
AW11	All Wales	344	163	4 weeks	Winter	Yes	10.1	1639
AW12	All Wales	344	163	4 weeks	Winter	No	10.1	1639
AW21	All Wales	344	169	4 weeks	Summer	Yes	11.1	1870
AW22	All Wales	344	169	4 weeks	Summer	No	11.1	1870
NE11	North East	38	16	4 weeks	Winter	Yes	6.2	99
NE12	North East	38	16	4 weeks	Winter	No	6.2	99
NE21	North East	38	19	4 weeks	Summer	Yes	4.2	79
NE22	North East	38	19	4 weeks	Summer	No	4.2	79
NW11	North West	51	21	4 weeks	Winter	Yes	10.5	221
NW12	North West	51	21	4 weeks	Winter	No	10.5	221
NW21	North West	51	30	4 weeks	Summer	Yes	10.6	319
NW22	North West	51	30	4 weeks	Summer	No	10.6	319
SE11	South East	205	96	4 weeks	Winter	Yes	11.6	1116
SE12	South East	205	96	4 weeks	Winter	No	11.6	1116
SE21	South East	205	95	4 weeks	Summer	Yes	12.7	1202
SE22	South East	205	95	4 weeks	Summer	No	12.7	1202
SW11	South West	51	24	4 weeks	Winter	Yes	9.4	226
SW12	South West	51	24	4 weeks	Winter	No	9.4	226
SW21	South West	51	25	4 weeks	Summer	Yes	6.8	171
SW22	South West	51	25	4 weeks	Summer	No	6.8	171

Additional Constraints

Thirdly, we consider how constrained the problem is, and for this we consider both the inclusion and exclusion of a specific constraint; constraints 3.17. Presently the Welsh Blood Service schedules clinics not only based on fulfilling demand but also based on satisfying clinic staff contractual hours. This leads to inefficiencies such as more clinics being scheduled than necessary to meet demand and potential overcollection and therefore potential wastage of blood products. Working with the clinic planner at the Welsh Blood Service, we formulated these constraints in the form of (3.17) whereby each region has a minimum number of clinics that must be scheduled per four-week planning horizon. We consider the exclusion of these constraints to measure their effect on the efficiency of the optimal clinic schedule.

Ideally, it would provide a more efficient collection model if constraints 3.11 could be removed. These constraints represent the obligation that the Welsh Blood Service have to the Welsh Government to ensure that even extremely rural and inefficient clinics are operated to provide those in the area with the opportunity to donate blood if they so wish. This means that each clinic must be scheduled at least once each calendar year.

Instance Nomenclature

In addition to the above variations, we also consider each of these base instances under each of the three objective functions, giving a total of 60 test instances. For ease of analysis and referral, each instance is given a name consisting of two letters followed by four e.g. SW1123. The instance nomenclature has been developed as follows and is explained using the same example;

- **Region:** The two letters at the start of each instance name represents the considered region e.g. in SW1123, the region is South West.
- **Planning Horizon:** The first number denotes the season; '1' represents the winter planning horizon and '2' represents the summer planning horizon.

Therefore, in the chosen example SW1123, the considered planning horizon is that of winter.

- **Additional Constraints:** The second number represents the inclusion (denoted by '1') or exclusion (denoted by '2') of the additional constraints, as discussed above e.g. in the case SW1123, the additional constraints are included.

- **Objective Function:** A final number may be added to represent the objective function used i.e. '1' denotes objective function one (the minimisation of clinics scheduled), '2' denotes objective function two (minimise estimated overcollection), and '3' denotes objective function three (the combination of both prior objective functions). In the given example, SW1123, objective three is considered. In absence of this third number, the more general base instance is being referred to.

5.7 Development of the Blood Collection Workforce Scheduling Model in Python

Similarly to the BDCSM in Section 5.6, an optimisation model for the scheduling of the clinic-based workforce at the WBS was created using Python and the linear programme package PuLP. This model follows the full formulation for the BCWSP described in Chapter 4. The BCWSM forms the second stage of the full scheduling model for the Welsh Blood Service.

5.7.1 Aims of Model

The BCWSM was built in Python to schedule the WBS clinic-based workforce optimally i.e. to optimally assign workers to the scheduled clinics in the output of stage one of the model, the BDCSM. The output of the stage two model is a workforce rota that aims to minimise workforce associated costs for the WBS.

The benefits of this model being used in place of the current practice at the WBS include a reduction in manual scheduling for the workforce planning team and an ability to increase cost efficiency. This is in answer to research question two: how can mathematical modelling help to schedule the clinic-based workforce at the Welsh Blood Service?

5.7.2 Development of Model

Similarly to the development of the BDCSM in Python, the BCWSM was developed incrementally; beginning with a small instance with only workers based in the South West region to schedule over one week, expanding to a model where all regions are able to be considered (one at a time) over a four-week period. This formed the original model for the BCWSP, described in Section 4.3. Following experimental trials of this model, it was evident that some solutions were poor, with unfair distribution of hours across workers. This prompted modifications to be made to the model to improve the viability of the solutions for the WBS, and formed the model described in Section 4.4.

In contrast to the stage one model, stage two only considers one region at a time; this is because the scheduling of the clinic-based workers in a given region is completely independent from that of other regions. The WBS only considers assigning clinic-based workers to a region different from their base region in cases such as a staff shortage due to sickness or emergency etc. These are therefore short-notice operational offline decisions and do not form part of the initial workforce scheduling process.

Input of Clinic Schedule and Clinic Data

The output of stage one of the model i.e. the BDCSM is a clinic schedule which is read into stage two of the model in the form of a dataframe using Pandas. The corresponding output of stage one is selected using a set of arguments relating to parameters of the model such as region and season.

From this schedule, the planning horizon is determined using the column headers. The clinic data is read into the model using the same data and method as described in Section 5.6.2 and a function is utilised to associate each scheduled clinic with its corresponding data. This is necessary for the model to obtain relevant information for workforce scheduling, such as the minimum number of workers per role required to operate a given clinic and the expected length of a working day for workers assigned to the clinic.

The workforce demands for a scheduled clinic consist of clinic roles and driving roles, as described in Section 4.1.2. A series of 'if' statements in the model main script sorts through each scheduled clinic and assigns it a 'resource demand matrix'; this contains the minimum number of each clinic role and driving role required at a given clinic.

Input of Workforce Data

Data concerning the clinic-based workforce that is relevant to the workforce scheduling process was collated into one Microsoft Excel file, separated into separate sheets by region. Since the workforce data provided by the WBS was anonymised, a unique staff number is generated to represent each clinic-based worker. The relevant data also includes, for each worker, their clinic role, any prospective driving roles, their number of weekly contracted hours, their base region, their working pattern (if any), their salary band and estimated hourly pay rate. This data was provided by the WBS with the exception of the estimated hourly pay; this was calculated using the salary range per salary band for the NHS which is openly accessible online. Since the pay rate for an NHS employee is dependent on the length of employment within the organisation, and the length of employment per worker was not able to be provided by the WBS, the mid-range salary per band is taken and divided by full-time hours to determine the hourly rate per salary band. This dataset is read into the model using the Python package, Pandas.

Worker Modes

Due to limitations of PuLP, in order to include working modes for each worker and to keep the problem linear, the worker modes were created independently from the worker variables.

A binary $|\mathcal{M}| \times 7$ 'mode matrix' is created for each worker, where $\mathcal{M}$ denotes the set of modes and each column represents a role; there are four columns for a driving role per vehicle (lorry, minibus, support vehicle and mobile donation unit) and three columns for the clinic roles (Registered Nurse, Supervisor and Clinic Care Assistant/Clinic Support Assistant). Each row of a mode matrix represents a potential working mode for a given worker, where any '1' values denote that the worker would perform the corresponding role of the given column if this mode is selected. A worker may work at most one driving role and one clinic role at any time. Examples of the mode matrices are presented in Tables 4.3 and 4.4.

These worker mode variables are included in the decision variable, where all modes are iterated over for each worker. Since these are to be iterated over independent of workers, all workers are required to have the same number of modes. To solve this issue, the parameter 'maximum modes' was introduced, which takes the value of the maximum number of unique modes of any worker considered in a given instance. To ensure that all workers have the same number of modes, duplicates of the final row of a given worker's mode matrix are appended to the matrix until the number of rows (and therefore modes) is equivalent to the maximum modes.

Worker Availability Patterns

Most clinic-based workers at the WBS are available to work any day that may be assigned to them (with the exception of any booked annual leave or sickness) but a small minority of workers have an agreed working pattern with the WBS. These specific working day patterns mean that they are not able to work on certain weekdays, and is often due to reasons like childcare or other commitments. To ensure that the model considers these working day patterns, we have created a function that requires a specific worker and the planning horizon as arguments, and returns the weekday availability of a worker. This is in the form of a list of binary values that correspond to each day of the planning horizon consecutively, where '0' denotes that they are unavailable and '1' denotes that they are available to be assigned to a clinic on the given day.

Annual Leave and Training

Since annual leave entitlement is dependent on the length of employment with the NHS (ranging from 27 to 33 days plus bank holidays), and we were not provided with this information by the WBS, we assume that each worker has a total of 30 days annual leave. Additionally, each worker is required to undertake a minimum of three training days per calendar year. To enable the model to consider annual leave and training to ensure that solutions are somewhat realistic, we randomly generate an annual leave and training schedule for the clinic-based workforce using a probability distribution relating to their 30 days of annual leave and three days of training over a calendar year. Two versions of this schedule are created, a descriptive version with entries consisting of either 'available','annual leave' or 'training'. Additionally, a binary version of this schedule is created with a '0' to denote any days that a worker is not available to be assigned to work a blood donation clinic i.e. they are scheduled for either a day of annual leave or training, and a '1' otherwise.

These schedules are read into the main script of the model and compared with worker weekday availability to ensure that no training days are scheduled for a worker on a day where they are unavailable to work. Any training days where this is the case are removed from the annual leave and training schedules. Each worker is assigned a dummy clinic corresponding to the correct number of hours for one day of annual leave, based on a worker's weekly contracted hours. This is the clinic that a worker will be scheduled to 'work' for any days of annual leave to ensure that these hours are also considered in the objective function. Similarly, there is one dummy training clinic for all workers since one day of training is always equivalent to 7.5 hours, regardless of contracted hours.

Sickness Leave

To provide the option for the model to consider sickness of workers, an argument for the model is created to determine if sick leave is included. If sick leave is to be included, the annual leave and training schedules are altered to randomly assign sick days in replacement of any days that are not already determined to be training or annual leave days. For the binary version of this schedule, any sick days are denoted with a '0' value to represent that a given worker is not available to work on a given day. These sick days are assigned according to a probability distribution with the assumption that sickness is at a rate of 6% for the clinic-based workforce; this figure was provided by the WBS as the mean sickness rate of 2019 for the clinic-based workforce.

Decision Variables, Objective Functions and Constraints

As described in the formulation of the BCWSP in Section 4.3, the model considers a total of five types of decision variable; $y_{s,m,i,t}$ (binary decision variable to decide if worker $s \in \mathcal{S}$ is assigned to work clinic $i \in \mathcal{I}$ on day $t \in \mathcal{T}$ in working mode $m \in \mathcal{M}$), $v_{1,s}$ (hours scheduled for worker $s \in \mathcal{S}$ in excess of full time hours), $v_{2,s}$ (number of overtime hours scheduled for worker $s \in \mathcal{S}$), $z^{\text{over}}_{s,w}$ and $z^{\text{under}}_{s,w}$ (number of hours scheduled for worker $s \in \mathcal{S}$ per week $w \in \mathcal{W}$ as overtime and 'undertime', respectively).

The model has two alternative objective functions; the minimisation of total overtime pay or the minimisation of total pay for scheduled hours, alongside the minimisation of weekly overtime and undertime per worker in the modified version of the model. Both of these objective functions also include a penalty function to ensure that Deputy Supervisors are only scheduled for the clinic role of Supervisor if there are no Supervisors available. The objective function to be considered is decided each time the model is run. All constraints described in Section 4.3 are included in the model, with additional constraints detailed in Section 4.4 also included in the modified version of the model.

Testing and Validation of Results

Throughout the incremental development of the stage two model, the solutions have been inspected to ensure that they are as expected. Similarly to stage one, to ensure that any solution reached by the model meets all of the included constraints, an assert statement is considered for each constraint to ensure that it holds true. This ensures that in circumstances where an optimality gap or time limit is considered during the 'solve' process, the solution is still valid.

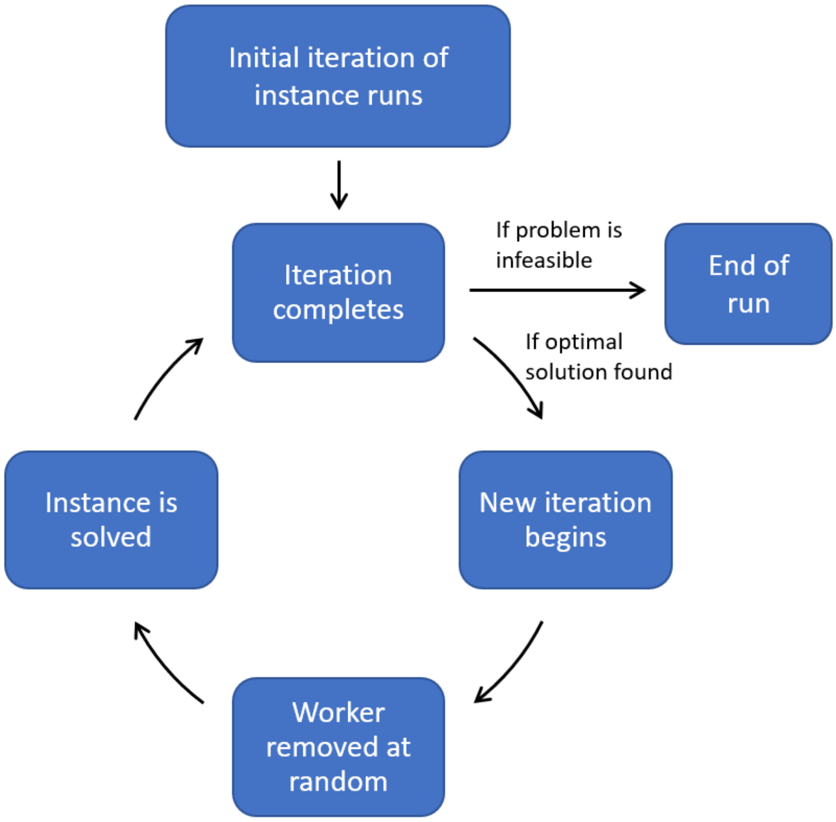
5.7.3 Development of Experimental Scenarios

To assess the solutions of the model, and how the model performs under different scenarios, we design various instances for experimental trials of the model. To consider workers in excess of the actual WBS clinic-based workforce, we create 'dummy' workers; these consist of a duplicate worker for each actual worker, but with a different unique staff number. This enables the effect of staffing level on the goodness of solutions to be observed, with the decision to remove a worker at random once each iteration of an instance is complete. Iterations of a given instance begin with double the actual WBS workforce for the given region (due to the dummy workers), and end once the problem is no longer feasible; this process is illustrated in Figure 5.5.

Test Instances

To assess the solutions of the model, and how the model performs given different scenarios, we considered various instances which can be found in Table 5.2. In addition to the three variables identified in stage one of the model i.e. the BDCSM, there were three main factors identified to likely have a significant effect on the

Figure 5.5: Worker Removal Process



goodness of a solution and the computation time: the prioritisation of the removal of dummy workers before any actual WBS workers, the removal of driving role constraints, and the inclusion of sickness leave.

Since objective function three for the BDCSM is concluded to be the preference of the alternative objective functions (an explanation for this is provided in Section 6.2.3), we only consider the output of these instances for stage two of the model; specifically, we utilise the output from the first iteration of each of these instances.

To illustrate the differences between the instances designed for stage two of the model, Table 5.2 displays this for the North East region (when solved independently of other regions in stage one of the model i.e. not an 'AW' instance). These instances are the same for all other regions with the letters at the beginning of an instance altered to represent the corresponding region of the instance. For cases where the region was solved collectively with other regions in stage one, an additional 'AW' is added before the specific region's letters.

Table 5.2: Stage Two - Instances for the North East Region Utilising Objective Function One

Instance Name	**Region**	**Number of Scheduled Clinic Days**	**Planning Horizon Length**	**Planning Horizon Season**	**Additional Constraints Included**	**Priority Removal of Dummy Workers**	**Driving Role Constraints Included**	**Sickness Leave Included**
NE1131111	North East	15	4 weeks	Winter	Yes	Yes	Yes	No
NE1131112	North East	15	4 weeks	Winter	Yes	Yes	Yes	Yes
NE1131121	North East	15	4 weeks	Winter	Yes	Yes	No	No
NE1131211	North East	15	4 weeks	Winter	Yes	No	Yes	No
NE1231111	North East	14	4 weeks	Winter	No	Yes	Yes	No
NE1231112	North East	14	4 weeks	Winter	No	Yes	Yes	Yes
NE1231121	North East	14	4 weeks	Winter	No	Yes	No	No
NE1231211	North East	14	4 weeks	Winter	No	No	Yes	No
NE2131111	North East	15	4 weeks	Summer	Yes	Yes	Yes	No
NE2131112	North East	15	4 weeks	Summer	Yes	Yes	Yes	Yes
NE2131121	North East	15	4 weeks	Summer	Yes	Yes	No	No
NE2131211	North East	15	4 weeks	Summer	Yes	No	Yes	No
NE2231111	North East	13	4 weeks	Summer	No	Yes	Yes	No
NE2231112	North East	13	4 weeks	Summer	No	Yes	Yes	Yes
NE2231121	North East	13	4 weeks	Summer	No	Yes	No	No
NE2231211	North East	13	4 weeks	Summer	No	No	Yes	No

Priority Removal of Dummy Workers

As portrayed in Figure 5.5, a worker is removed at each iteration of the model for a given instance. The option is provided by the model to prioritise dummy workers to be removed before any actual WBS workers. This helps to avoid any potential infeasibility that could be caused by removing workers completely at random, as there may be insufficient numbers of a specific role. This also enables the observation of how the model performs with actual staff (once all dummy workers are removed) as the staffing level decreases.

The motivation behind including the option to remove a worker completely at random (regardless of if a worker is a dummy worker or not) gives the potential to gain insight into the balance of skills and/or roles and if this could be improved. For example, as presented in Table 5.3, the numbers of workers in the role of Registered Nurse (RN) is limited in several regions, while this number appears inflated for the South East region.

Driving Role Constraints

The model also provides the option to exclude driving role demand constraints and to exclude driving roles from mode matrices. This allows the observation of how far the iteration over driving role modes increases the computation time. Additionally, this enables us to gain insight into how much the problem is constrained by the limited number of drivers per region, as displayed in Table 5.4. By removing the constraints that require the driving role demands to be met per scheduled clinic, in addition to removing driving roles from each workers mode matrix, this is effectively simulating the scenario where all clinic-based workers are able to drive any vehicle.

Sickness Leave

Though sickness leave is usually determined with short-notice to the planning team at the WBS, with operational offline decisions needing to be made to counteract

Table 5.3: Number of Workers per Region

	Region			
	North East	North West	South East	South West
Number of Workers Considered in the First Instance	34	26	154	36
Number of Actual WBS Workers	17	13	77	18
Number of Actual WBS RNs	3	2	16	3
Number of Actual WBS Supervisors	1	1	4	1
Number of Actual WBS Deputy Supervisors	1	1	4	1
Number of Actual WBS CCA/CSAs	12	9	53	13

RN = Registered Nurse; CCA = Clinic Care Assistant;
CSA = Clinic Support Assistant

Table 5.4: Number of Drivers per Vehicle per Region

	Region			
	North East	North West	South East	South West
Number of Actual WBS Lorry Drivers	3	3	7	5
Number of Actual WBS Minibus Drivers	3	3	7	5
Number of Actual WBS SV Drivers	0	0	8	0
Number of Actual WBS MDU Drivers	0	3	0	0

SV = Support vehicle; MDU = Mobile donation unit any negative impact of sickness on the operation of clinics, considering an estimated sickness rate in the model allows us to observe this impact on the staffing level requirements.

Instance Nomenclature

Table 5.2 display the instances for the North East region (when stage one is solved independent of other regions) when objective function one is considered. Each of these instances are also included where objective function two is considered, giving a total of 256 initial test instances (32 for each region when solved either collectively or independently of other regions in stage one, for each objective function). For ease of analysis and referral, each instance is given a name building on the names given in stage one of the model. The instance nomenclature has been developed as follows and is explained using the same example;

- **Region:** The two letters at the start of each instance name represents the considered region e.g. in SW1131112, the region is South West.
- **Planning Horizon:** The first number denotes the season; '1' represents the winter planning horizon and '2' represents the summer planning horizon. Therefore, in the chosen example SW1131112, the considered planning horizon is that of winter.
- **Additional Constraints:** The second number represents the inclusion (denoted by '1') or exclusion (denoted by '2') of the additional constraints in stage one of the model e.g. in the case SW1131112 the additional constraints are included.
- **Stage One Objective Function:** Following this, since we only consider output from stage one of the model that utilises objective function three, a '3' is always in this position for all stage two instances.
- **Stage Two Objective Function:** This fourth number represents the objective function considered for the BCWSM; a '1' denotes objective function

one (the minimisation of total overtime pay, weekly overtime and weekly undertime) while a '2' denotes objective function two (minimisation of total pay for scheduled hours, weekly overtime and weekly undertime). In the given example, SW1131112, objective function one is used.

- **Priority Removal of Dummy Workers:**The fifth number denotes whether the removal of dummy workers is prioritised over the removal of any actual WBS workers, with a '1' representing cases where dummy workers are removed first, and a '2' represents cases where the removal of a worker is random, and not dependent on whether they are a dummy or actual worker. In the example of instance SW1131112, dummy worker removal is prioritised.

- **Driving Role Constraints:**The sixth number in the instance names for stage two represent whether or not driving role demand constraints are included, where a '1' denotes the inclusion of driving role constraints, and a '2' denotes the exclusion of these constraints. In the example instance of SW1131112, driving role constraints are included.

- **Sickness Leave:** Lastly, the final number in an instance name for stage two signifies whether sickness of workers is considered, where a '1' denotes cases where sickness is not considered, and a '2' denotes cases where sickness of workers is included, e.g. the instance SW1131112 does consider sickness leave.

5.7.4 Summary

In this chapter, the prototype tool for the Blood Donation Clinic Scheduling Problem (BDCSP) has been presented followed by the Blood Donation Clinic Scheduling Model (BDCSM) and Blood Collection Workforce Scheduling Model (BCWSM) both developed in Python, using PuLP as a linear programme modeller. For each model, the aims of the model and the use and format of data was described in addition to how the model was developed. For both of the models developed in Python, various test instances were presented alongside the motivation behind the design of the instances, and the instance nomenclature has been described in detail.

In the following chapter, Chapter 6, the results of these experimental test instances are presented with an analysis of the insights gained from the solutions and how the current practice of both clinic and workforce scheduling at the Welsh Blood Service could be more efficient.

Chapter 6

Experimental Results

In this chapter, the results of the experimental scenarios for the Blood Donation Clinic Scheduling Model (BDCSM) and the Blood Collection Workforce Scheduling Model (BCWSM) are presented. This chapter is structured as follows: Section 6.1 describes the hardware utilised to run the experimental scenarios for both stages of the scheduling model (the BDCSM and the BCWSM) in addition to the software used. Section 6.2 presents the computational results of the BDCSM and discusses the solutions of the selected test instances. Similarly, Section 6.3 presents the computational results of the BCWSM alongside the solutions of the chosen test instances. The two latter sections each close with insights gained from the experimental results.

6.1 Equipment

All test instances for both the BDCSM and the BCWSM were run remotely on servers (located in the university) limiting each instance to run on only one core utilising GNU Parallel software [94]. Since PuLP and COIN-OR CBC solver are limited to only utilising one core during the preprocessing phase, the ability to utilise more than one core is restricted. This also recreates similar circumstances to how the model might be run on site at the Welsh Blood Service (WBS) if it were implemented, since they have a server estate in their data centre where their current digital rostering system, for example, is hosted and run remotely from PCs.

6.1.1 Hardware

All test instances (unless stated otherwise) were run remotely on servers with a total of 64 cores. The server workstation consists of four processors (AMD Opteron 6366HE - 1.8GHz, 16Mb cache) with each processor having 16 cores. The memory of the workstation is 8 $\times$ 32GB 1600MHz DDR3 ECC Reg with Parity DIMM Dual Rank.

6.1.2 Software

Both of the models developed in Python (BDCSM and BCWSM) utilise Python version 3.8.3 and PuLP version 2.3 which uses COIN-OR CBC version 2.9. Other Python packages are used such as Pandas (version 1.0.5), NumPy (version 1.18.5) and GeoPy (version 2.0.0). A virtual environment was created for each of the two models to certify that identical versions of all software packages are utilised by the models for all results; this ensures that all results are reliable and that the code functions as expected. Additionally, GNU Parallel 2018 was utilised to run each test instance on exactly one core.

6.2 Blood Donation Clinic Scheduling Model

The Blood Donation Clinic Scheduling Model developed in Python, detailed in Chapter 5, was run for each problem instance described in Table 5.1 for each of the three alternative objective functions. The model was run with an optimality gap of 0.5% i.e. a solution may be considered optimal if it is within 0.5% of the best known solution; this is to enable a solution to be reached within a smaller time. The experimental results are presented in this section, and how the model can increase efficiency of the clinic scheduling process for the WBS, in answer to research question one (described in Section 1.5).

6.2.1 Computational Results

To ensure the computational results are reliable, we ran each instance 10 times and collated the results. The following tables (categorised by objective function) detail the median, minimum, and maximum solution times and total run times for each instance. As a reminder to the reader, the instance nomenclature for the BDCSM follows the structure of A/B/C/D where:

- A denotes the region considered i.e. {AW = All Wales; NE = North East; NW = North West; SE = South East; SW = South West}
- B denotes the planning horizon i.e. {1 = winter; 2 = summer}
- C denotes the inclusion of the additional constraints i.e. {1 = minimum number of clinics to be scheduled per region included; 2 = minimum number of clinics to be scheduled per region excluded}
- D denotes the objective function i.e. {1 = minimisation of clinic days scheduled; 2 = minimisation of estimated overcollection; 3 = minimisation of both clinic days scheduled and estimated overcollection}

Tables 6.1, 6.2 and 6.3 also exhibit the number of decision variables and the number of constraints for each instance. In the instances for the model where all regions are solved collectively i.e. AW instances, the total number of constraints considered reaches a maximum of over 2.2 million. Despite this, the maximum solution time for these large instances is 526 seconds (AW212) with a maximum total run time of around eight hours for all regions solved collectively (AW213). In comparison to the current practice of manually scheduling clinics over four-week periods, the model performs well, providing an optimal (or near optimal) clinic schedule in a much shorter time frame.

Table 6.1: Objective Function 1 - Computational Results

Instance	**Number of Decision Variables**	**Number of Constraints**	**Median Solution Time (secs)**	**Min. Solution Time (secs)**	**Max. Solution Time (secs)**	**Median Total Run ime (secs)**	**Min. Total Run Time (secs)**	**Max. Total Run Time (secs)**
AW111	10,000	2,263,681	101.4	89.8	121.1	4,599.3	4,190.3	5,401.6
AW121	10,000	2,263,677	85.6	79.4	94.2	4,177.2	3,810.4	4,497.4
AW211	10,000	2,263,681	83.5	77.8	87.4	4,180.3	3,996.3	4,494.4
AW221	10,000	2,263,677	88.9	81.9	101.3	4,296.1	3,937.3	5,021.5
NE111	1,432	212,241	6.5	5.8	7.0	108.4	95.0	116.7
NE121	1,432	212,240	8.0	7.6	10.5	103.1	100.2	131.9
NE211	1,432	212,241	6.6	5.8	7.1	102.0	94.9	111.9
NE221	1,432	212,240	9.5	9.1	9.9	103.7	102.9	104.7
NW111	1,796	137,202	4.8	4.4	5.2	134.1	129.9	136.2
NW121	1,796	137,201	5.8	5.5	6.9	120.5	118.0	157.5
NW211	1,796	137,201	4.7	4.3	5.5	133.1	130.1	175.3
NW221	1,796	137,200	5.4	5.2	6.0	119.0	116.8	120.6
SE111	6,108	1,633,127	57.7	52.8	64.7	1,425.1	1,391.5	1,645.5
SE121	6,108	1,633,126	68.2	64.4	70.7	1,334.6	1,290.6	1,408.0
SE211	6,108	1,633,128	57.4	56.6	64.9	1,598.5	1,573.3	1,641.8
SE221	6,108	1,633,127	68.8	65.9	74.6	1,490.3	1,450.7	1,543.4
SW111	1,796	133,847	11.9	9.6	13.4	1,202.2	1,085.3	1,316.5
SW121	1,796	133,846	13.2	11.2	15.1	782.6	701.1	846.7
SW211	1,796	133,847	7.7	7.2	8.1	862.0	801.3	894.9
SW221	1,796	133,846	14.8	14.2	16.6	843.9	745.7	852.5

Table 6.2: Objective Function 2 - Computational Results

Instance	**Number of Decision Variables**	**Number of Constraints**	**Median Solution Time (secs)**	**Min. Solution Time (secs)**	**Max. Solution Time (secs)**	**Median Total Run Time (secs)**	**Min. Total Run Time (secs)**	**Max. Total Run Time (secs)**
AW112	10,000	2,263,681	150.6	132.9	168.0	1,890.9	1,582.0	1,966.9
AW122	10,000	2,263,677	198.6	181.8	204.9	1,799.4	1,675.0	1,842.3
AW212	10,000	2,263,681	486.3	444.8	526.1	2,216.2	1,986.7	2,288.5
AW222	10,000	2,263,677	128.7	120.5	140.6	1,644.6	1,499.0	1,682.1
NE112	1,432	212,241	10.1	9.8	10.9	80.7	79.4	81.3
NE122	1,432	212,240	25.9	24.5	28.1	86.0	85.1	88.8
NE212	1,432	212,241	48.5	44.5	52.8	118.4	114.8	124.0
NE222	1,432	212240	61.6	54.2	69.2	134.4	108.2	151.7
NW112	1,796	137,202	960.5	650.5	1,082.0	1,033.3	721.4	1,157.6
NW122	1,796	137,201	29.4	27.9	32.1	92.7	90.8	102.7
NW212	1,796	137,201	7.1	6.7	7.6	82.5	72.4	83.2
NW222	1,796	137,200	6.0	5.3	6.3	69.2	67.2	73.5
SE112	6,108	1,633,127	216.4	209.5	220.9	988.6	981.5	1,008.3
SE122	6,108	1,633,126	178.6	167.1	188.1	773.3	729.3	816.2
SE212	6,108	1,633,128	139.9	120.1	151.8	994.4	881.0	1,098.8
SE222	6,108	1,633,127	125.0	117.2	159.7	691.5	658.2	945.5
SW112	1,796	133,847	235.3	221.8	266.8	1,117.9	1,044.2	1,366.7
SW122	1,796	133,846	342.5	266.6	403.9	1,085.6	979.1	1,186.5
SW212	1,796	133,847	24.9	22.0	26.2	783.8	695.9	820.5
SW222	1,796	133,846	273.8	216.8	395.4	1,101.4	963.2	1,329.0

Table 6.3: Objective Function 3 - Computational Results

Instance	Number of Decision Variables	Number of Constraints	Median Solution Time (secs)	Min. Solution Time (secs)	Max. Solution Time (secs)	Median Total Run Time (secs)	Min. Total Run Time (secs)	Max. Total Run Time (secs)
AW113	10,000	2,263,681	79.5	68.4	92.9	17,736	14,124	19,554
AW123	10,000	2,263,677	70.5	65.1	83.5	22,452	18,930	25,914
AW213	10,000	2,263,681	78.1	67.6	115.5	25,278	22,332	28,380
AW223	10,000	2,263,677	74.4	64.0	86.6	13,926	12,270	15,948
NE113	1,432	212,241	1.9	1.8	1.9	474	444	498
NE123	1,432	212,240	2.1	1.8	2.6	1,242	1,062	1,380
NE213	1,432	212,241	2.0	2.0	2.0	1,110	1,062	1,230
NE223	1,432	212,240	2.3	1.8	2.7	1,698	1,446	1,986
NW113	1,796	137,202	4.0	3.9	4.4	7,998	7,596	8,988
NW123	1,796	137,201	2.3	2.3	2.4	1,638	1,554	1,950
NW213	1,796	137,201	2.4	2.3	2.4	840	810	870
NW223	1,796	137,200	2.0	2.0	2.1	408	390	450
SE113	6,108	1,633,127	33.8	31.7	35.4	29,724	26,040	32,532
SE123	6,108	1,633,126	29.1	25.6	29.5	18,954	17,364	19,956
SE213	6,108	1,633,128	27.7	26.6	32.1	9,678	9,330	10,524
SE223	6,108	1,633,127	32.0	29.2	33.9	31,758	29,820	34,314
SW113	1,796	133,847	17.1	15.8	17.8	6,708	6,150	7,524
SW123	1,796	133,846	20.4	18.7	22.7	14,328	11,958	16,248
SW213	1,796	133,847	20.0	19.5	23.4	23,490	21,804	33,642
SW223	1,796	133,846	18.9	16.7	22.2	10,944	10,212	12,786

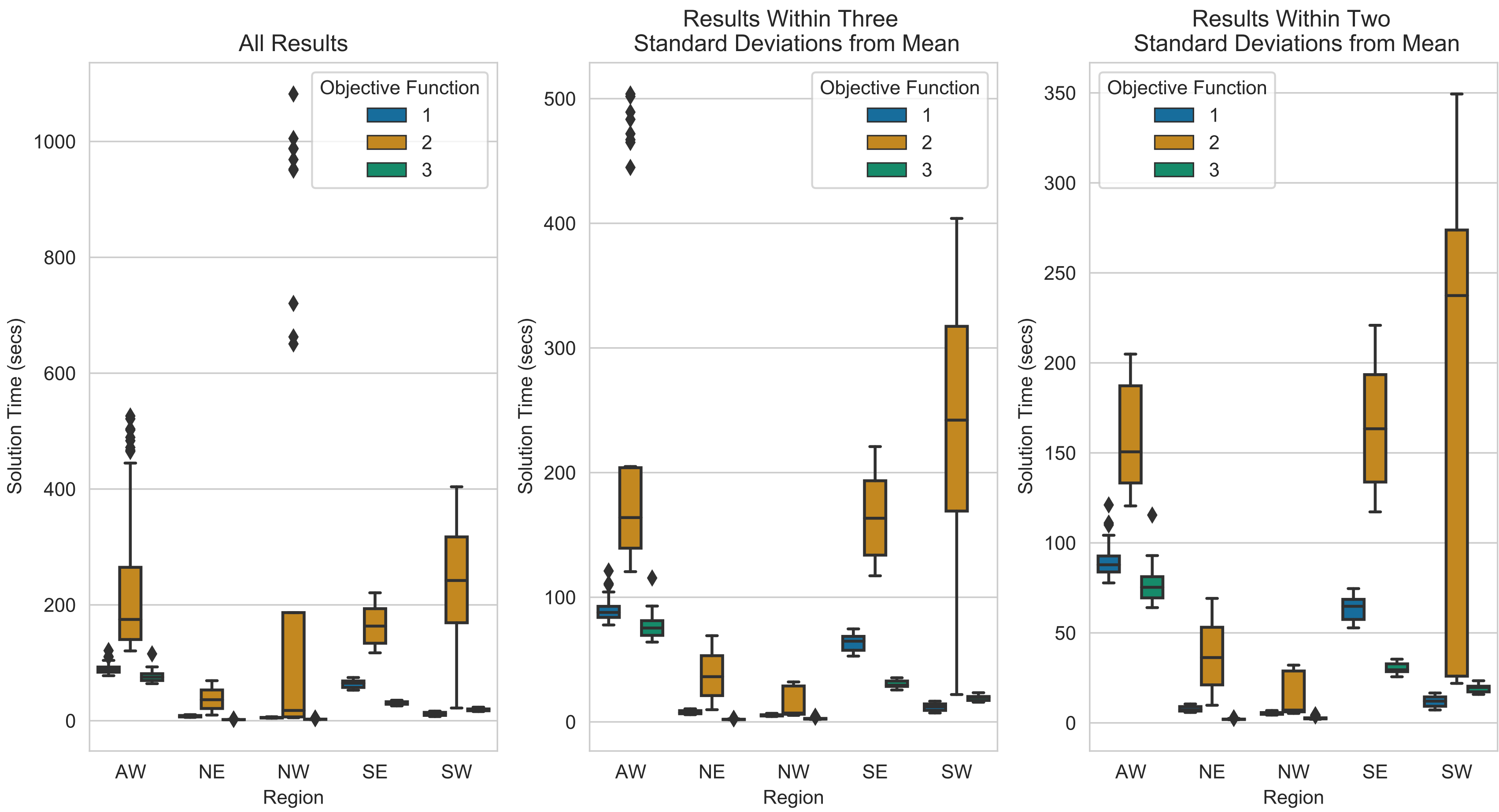
The number of constraints for each region does not follow the same increase as the number of decision variables, and this is due to a particular set of constraints. Constraints (3.21) ensure that any public clinics within a given number of miles of another public clinic should not be scheduled on the same day. Of the smaller regions (North East, North West and South West), the North East region has significantly more public clinics that are within five miles of another public clinic than the other smaller regions, with a total of 107 pairs of conflicting clinics. Following this is the North West region with 52 pairs of conflicting clinics and lastly the South West region with 45 pairs of conflicting clinics. This explains the trend in increase in number of constraints amongst these smaller regions.

Objective Function

Recall in Section 3.3.2, the following three alternative objective functions were introduced; objective function one is the minimisation of the number of clinic days scheduled (3.3), objective function two is the minimisation of estimated overcollection (3.4), whilst objective function three is the minimisation of both the number of clinic days scheduled and estimated overcollection (3.5). The effect of these objective functions on solution time is presented in Figure 6.1. The results are shown in full, followed by two graphs with the more extreme results excluded to increase the visibility for shorter solution times. Most of these extreme results are due to specific instances having longer solution times than others, e.g. instances NW112 and AW212 have much higher solution times than all other instances for these regions.

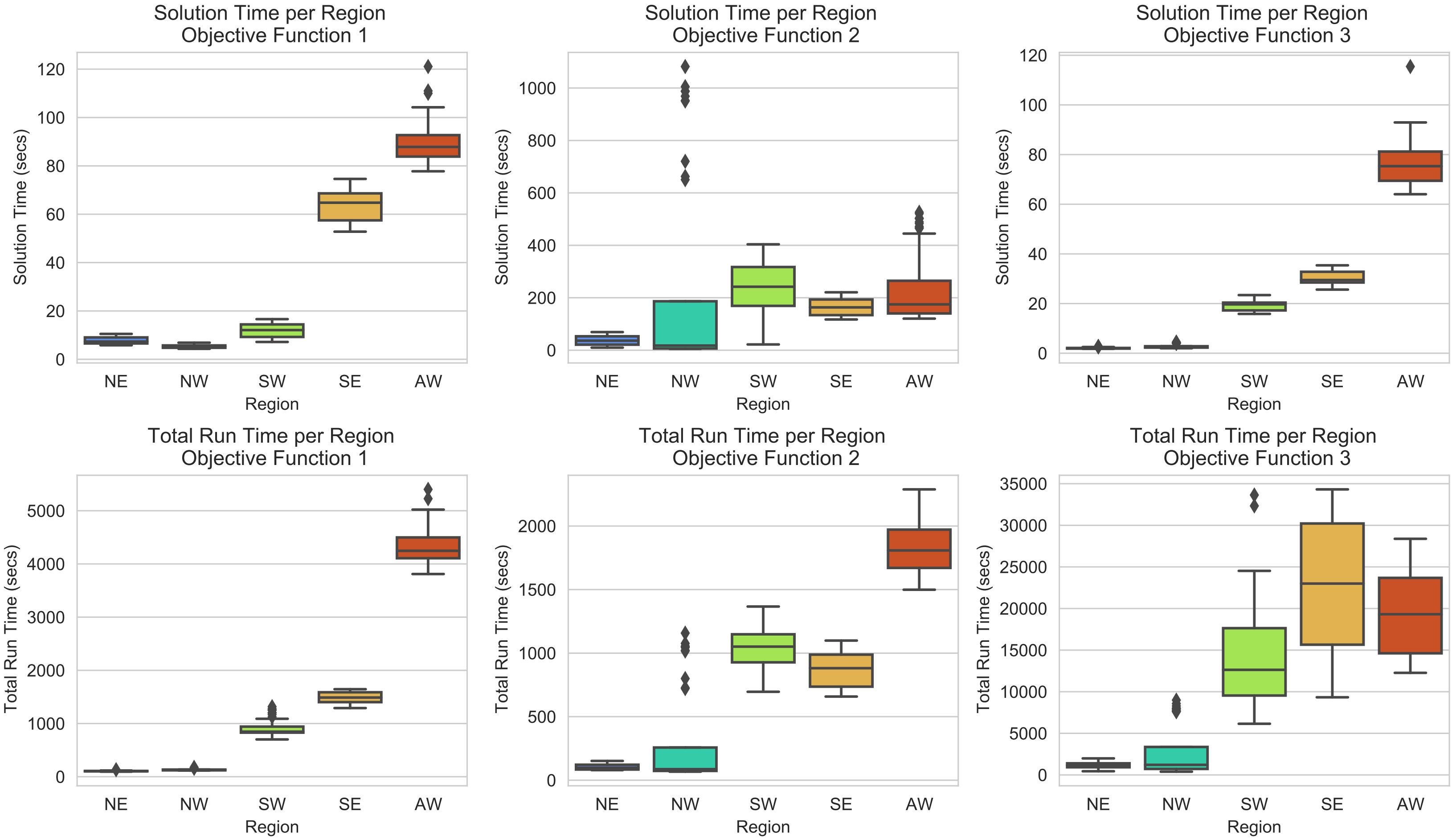
There is a noticeable increase in solution time for objective function two when compared to the other two alternative objective functions, across all regions. This is likely caused by the complexity associated with meeting demand exactly (i.e. minimisation of overcollection). The solution time for objective function three, despite it being the combination of both objective functions one and two, is the fastest solution of all alternative objective functions with the exception of the South West region, where objective function one marginally outperforms objective function three.

Figure 6.1: Solution Time per Region per Objective Function



Please note the box plots are in ascending order of objective function number for each region, for cases where the values are close to zero and the colour is not visible.

Figure 6.2: Solution Time and Total Run Time per Region per Objective Function



Regions along the x-axis are in ascending order of number of decision variables.

The total run time for objective function three is drastically slower than for objective functions one and two; this is likely due to the time taken to set up the objective function and all associated decision variables, as this objective function considers both number of clinics scheduled and weekly overcollection. However, the solution times for this objective function are all less than 116 seconds, and shows that the additional time in the total run time is caused by the time to set up the model.

Regions

The general trend across regions regarding computation times can be seen in Figure 6.2. This trend is as the number of decision variables increases, both the solution time and the total run time increase. However, this is not the case for instances that consider objective function two, most likely due to the complexity associated with meeting demand exactly i.e. achieving zero estimated overcollection.

Table 6.4: Mean Total Run Times for All Wales - Collective Model Vs. Independent Model

	Mean Total Run Time (secs)	
Instance	**Collective Model**	**Independent Model**
AW111	4,688.29	2,926.06
AW112	1,819.63	3,171.67
AW113	17,183.11	45,043.67
AW121	4,161.68	2,349.77
AW122	1,794.48	2,038.38
AW123	22,066.74	36,112.89
AW211	4,201.02	2,700.75
AW212	2,171.93	1,953.73
AW213	25,690.53	36,922.69
AW221	4,331.34	2,558.37
AW222	1,618.32	2,043.67
AW223	13,904.50	45,021.63

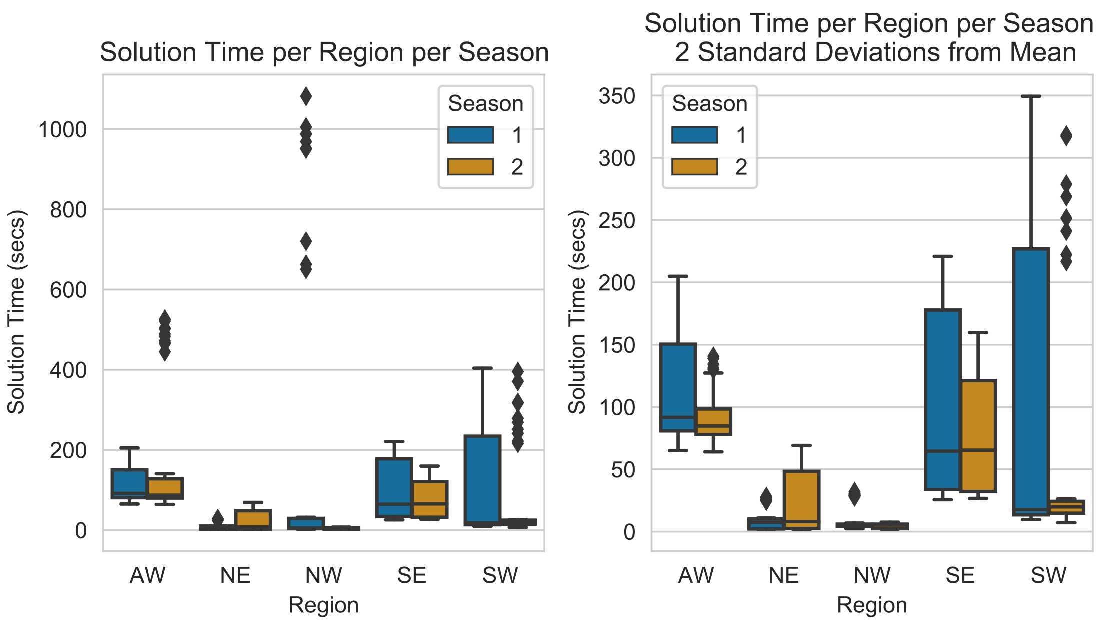
Table 6.4 presents the mean total run time per instance for the whole of Wales i.e. 'AW', both when the model solves for all regions collectively and when regions are solved independently, with the latter run times being the sum of the mean total run time per instance per region. For the cases where objective function three is considered, it is faster to solve the model for all regions collectively than

independently. This also applies to most instances considering objective function two, while the converse is true for objective function one.

Seasonality

The model was tested on two distinct four-week planning horizons to observe any potential seasonal trends and any effect this may have on complexity. Planning horizon one (or season one) is in the winter, whilst planning horizon two (or season two) is in the summer. There are differing trends in the relationship between season and solution time between the regions, as displayed in Figure 6.3. All of the solution times in this figure exceeding 116 seconds are from objective function two, with little contrast in solution time between seasons for objective functions one and three.

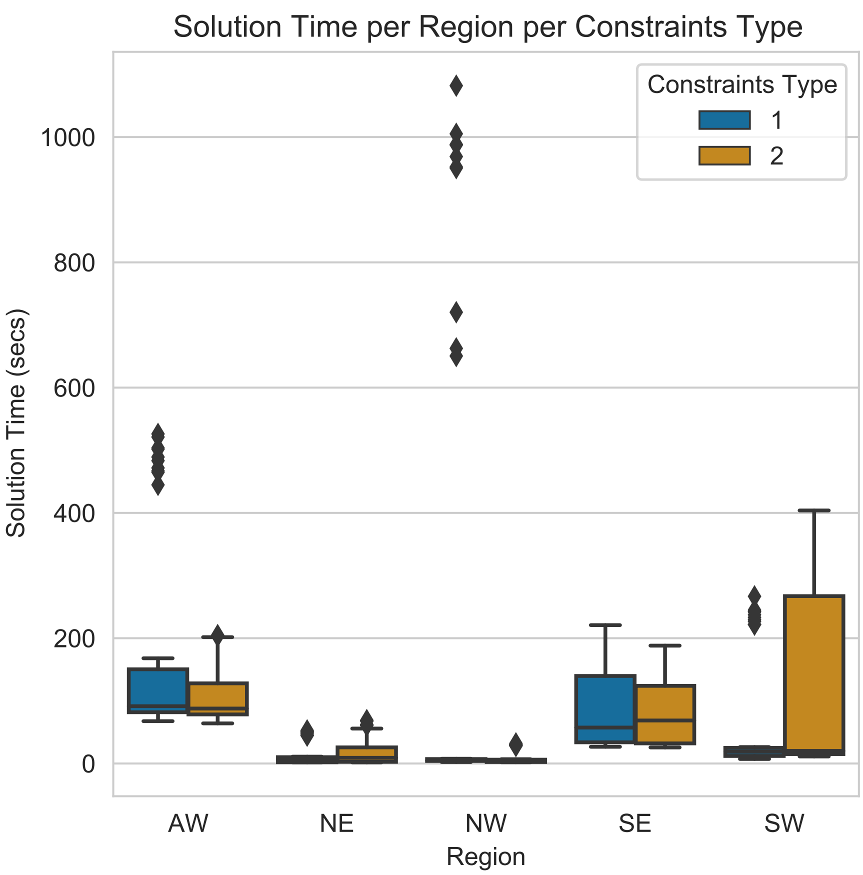
Figure 6.3: Solution Time per Region per Season



For North East instances, the season two instances have a greater solution time than that of season one instances. This is likely due to there being an average time window size per available clinic of 4.2 days for the region during the season two planning horizon, compared to 6.2 days for the season one planning horizon. In addition, the total available clinic days for season one for the North East region is 99 with fewer available clinic days in season two at 79.

Contrastingly, season one generally appears to have a greater solution time for most cases in all other regions, with the exception of some instances for 'AW' (where all regions are solved collectively) with the slower solve time likely associated with the model prioritising meeting demand exactly with less flexibility (discussed further in Section 6.2.1). For the North West and South East regions, the solution time is generally slower for season one than for season two, and this is likely due to these regions having a higher number of total available clinic days per planning horizon for season two (as displayed in Table 5.1). Finally, the South West region is generally solved in a similar time for season two and season one (according to the mean), with the exception of the SW212 instances being solved much faster than the other South West instances solved under objective function two. This is unsurprising as there are 24 clinics available to be scheduled in the South West region in season one and 25 in season two.

Figure 6.4: Solution Time per Region per Constraints Type



Additional Constraints

Alternative scenarios are considered for the constraints of the model where a set of constraints are either included ('constraints type one') or excluded ('constraints type two') for different instances. Figure 6.4 displays the solution time per region per constraints type, and illustrates that the instances with the slowest solution times consider constraints type one; this is the case for regions 'AW' and 'NW', more specifically, for the instances AW212 and NW112. The latter instance likely takes significantly slower to solve due to objective function two attempting to meet demand exactly (by minimising estimated overcollection) with less flexibility due to a minimimum of 15 clinics required to be scheduled in the North West region over the planning horizon by the included set of constraints. In addition to this, the North West region when solved independently has the lowest collection demand of all the regions, increasing difficulty to meet this low demand whilst scheduling the minimum number of clinics. Similarly for the AW212 instance, though regions are solved collectively and therefore do not have a pre-determined demand, the inclusion of the set of constraints that ensure a minimum number of clinics are scheduled per region gives the model significantly less flexibility to meet the demand exactly, and increases complexity.

Contrastingly, the South West region has a long solution time for constraints type two. This is caused by one particular base instance (with constraints type one) for the South West region under both objective functions one and two having a significantly faster solution time than the other base instances i.e. the apparent increase in solution time for constraints type two instead originates from the corresponding instance with constraints type one (SW212) having an unusually low solve time (with both objective functions one and two, as all South West instances under objective function three have a low solve time). This difference in computational time between instances SW212 and SW222 is likely due to the tighter average time window size per available clinic for the former at 6.8 days, compared to 9.4

days for the latter, contributing to fewer possibilities for the model to consider for SW212. This reflects findings in the literature regarding solution times and complexity, with Koné et al. [54] describing the 'easy-hard-easy' pattern associated with how constrained a scheduling problem is (in the form of a linear programme). With a relatively unconstrained problem, it is easy to solve, but the complexity increases as the problem becomes more constrained until it decreases again as there are few possible solutions remaining.

6.2.2 Schedule Solutions

The solutions of the Blood Donation Clinic Scheduling Model for all test instances are now discussed, to observe how well the model performs at providing a feasible clinic schedule for the Welsh Blood Service. Please note that when we refer to the number of clinics scheduled, this may include multiday clinics where each day of the multiday clinic is considered to be one clinic i.e. the total number of clinics scheduled is referring to the total number of clinic days scheduled.

Since dummy clinics were included in the model to support to assignment of workers to annual leave and training in the workforce scheduling model, this totals to an additional 240 clinic days scheduled each four-week planning horizon. When the number of clinic days scheduled are discussed in this section, the number consists of standard clinics only, with dummy clinics deducted. However, these 240 dummy clinic days are likely to play a role in the optimality gap of a given objective function value.

Objective Function

With three alternative objective functions tested for each instance, Table 6.5 displays the performance of the BDCSM under each objective. The performance can be judged on the two values that are considered in the objective functions (combined) i.e. total number of clinic days scheduled over the four-week planning horizon and total estimated units of blood collected in excess of the demand.

The values in Table 6.5 are the mode number of clinics scheduled and mode estimated overcollection. In the majority of cases, these values are the same across all iterations of each instance, however in three cases, there is some variation between iterations. Namely, the mode number of clinics scheduled for instance SW213 is 15, with mode overcollection being 1, however in some iterations these values are instead 16 and 0, respectively. This is due to the total objective value being the same in both cases as weights for each term in objective function three have not been introduced, and would need to be determined by the WBS to reflect their preferences and priorities. Similarly, the mode number of clinics scheduled in instance SW222 is 15, but in some iterations this value is 16. This is a result of objective function two minimising only the estimated overcollection, and thus the number of clinics scheduled is not considered in this objective function.

In most instances in Table 6.5, the values for both number of clinics scheduled and estimated overcollection for objective function three are the same or marginally higher than for the respective considered value in objective function one and objective function two. In some cases, objective function three even improves upon the respective considered value in the alternative objective functions; this is due to the optimality gap of the model. Each of these instances where either number of clinics scheduled or estimated overcollection is higher for objective function one or objective function two, respectively, than for objective function three, are denoted by a double asterisk in Table 6.5. Table 6.8 displays the results for each of these instances when the same model is run without an optimality gap. It is evident from these results that the optimality gap of 0.5% in the original results is responsible for the suboptimal values for these instances, as the results in Table 6.8 match the corresponding value for the same instance using objective function three.

Table 6.5: Solutions per Region for all Objective Functions (Solved Independently and All Wales Collectively)

	Objective Function 1		**Objective Function 2**		**Objective Function 3**	
Instance	**Number of Clinics Scheduled**	**Estimated Overcollection**	**Number of Clinics Scheduled**	**Estimated Overcollection**	**Number of Clinics Scheduled**	**Estimated Overcollection**
AW11	109	199	110	0	109	0
AW12	95	49	107	0	96	0
AW21	109	379	111	0	109	0
AW22	**82	21	88	0	81	1
NE11	15	97	15	26	15	26
NE12	14	59	14	8	14	8
NE21	15	71	15	15	15	15
NE22	13	34	14	**16	13	5
NW11	15	86	15	0	15	1
NW12	11	12	14	0	12	0
NW21	15	28	15	0	15	0
NW22	9	11	12	0	9	0
SE11	**65	67	67	0	64	0
SE12	63	20	67	0	64	0
SE21	64	194	65	0	64	0
SE22	53	83	59	0	53	0
SW11	**16	240	16	0	15	1
SW12	15	195	16	0	16	0
SW21	16	136	15	0	*15	*1
SW22	**15	185	*15	0	14	0

*Figures varied between two values over the total ten iterations of the specific instance.
**Values for either number of clinics scheduled or estimated overcollection are higher than expected when compared with the same instance under a different objective function.

Table 6.6: Number of Clinics Scheduled and Estimated Supply per Region for all Objective Functions (Solved Independently)

	Objective Function One		**Objective Function Two**		**Objective Function Three**	
Instance	**Number of Clinics Scheduled**	**Estimated Supply**	**Number of Clinics Scheduled**	**Estimated Supply**	**Number of Clinics Scheduled**	**Estimated Supply**
NE11	15	861	15	790	15	790
NE12	14	823	14	772	14	772
NE21	15	790	15	734	15	734
NE22	13	753	14	735	13	724
NW11	15	556	15	470	15	471
NW12	11	482	14	470	12	470
NW21	15	470	15	442	15	442
NW22	9	453	12	442	9	442
SE11	65	4,341	67	4,274	64	4,274
SE12	63	4,294	67	4,274	64	4,274
SE21	64	4,216	65	4,022	64	4,022
SE22	53	4,105	59	4,022	53	4,022
SW11	16	1,734	16	1,494	15	1,495
SW12	15	1,689	16	1,494	16	1,494
SW21	16	1,542	15	1,406	*15	*1,407
SW22	15	1,591	*15	1,406	14	1,406

*Figures varied between two values over the total ten iterations of the specific instance.

Table 6.7: Number of Clinics Scheduled and Estimated Supply per Region for all Objective Functions (Solved Collectively)

		Objective Function One		**Objective Function Two**		**Objective Function Three**	
Instance	**Region**	**Number of Clinics Scheduled**	**Estimated Supply**	**Number of Clinics Scheduled**	**Estimated Supply**	**Number of Clinics Scheduled**	**Estimated Supply**
AW11	NE	15	839	15	846	15	798
AW12	NE	6	411	11	620	8	514
AW21	NE	15	787	15	734	15	767
AW22	NE	3	208	6	356	6	402
AW11	NW	15	538	15	560	15	582
AW12	NW	5	266	13	497	5	232
AW21	NW	15	421	16	503	15	468
AW22	NW	1	15	5	231	1	64
AW11	SE	64	4,199	65	4,158	64	4,093
AW12	SE	68	4,605	67	4,308	67	4,514
AW21	SE	64	4,450	65	3,940	64	4,062
AW22	SE	62	4,696	62	4,438	58	4,439
AW11	SW	15	1,625	15	1,438	15	1,529
AW12	SW	16	1,769	16	1,577	16	1,742
AW21	SW	15	1,309	15	1,411	15	1,291
AW22	SW	16	1,690	15	1,563	16	1,684

Table 6.8: Number of Clinics Scheduled and Estimated Overcollection without Optimality Gap for Selected Instances

Instance	**Number of Clinics Scheduled**	**Estimated Overcollection**
AW221	81	17
NE222	13	3
SE111	64	52
SW111	15	159
SW221	14	50

All results obtained from running the model without an optimality gap on a PC with processor Intel(R) Core(TM) i7-8550U CPU @ 1.80GHz 2.00 GHz 16GB RAM.

The results displayed in Table 6.8 were obtained from running the model on a PC with a higher specification than the processors that the original results were run on. This caused the computation times of the model to decrease for all five instances when compared with the original computational results. This improvement in computation time ranges from a scale of two to three times faster than the original computational times for both the solve time and the total run time.

The number of clinics scheduled over a four-week planning horizon per region are displayed in Figure 6.5, with a comparison between the values for each objective function. The results when each region is solved independently are shown alongside the results for each region when all regions are solved collectively i.e. the AW instances. Since a predetermined percentage of total demand is allocated to each region when they are solved independently, this causes the lowest number of clinics scheduled to generally be higher than for the same region in AW instances in order to meet the demand. Figure 6.5 shows that generally, objective function three performs similarly to objective function one regarding number of clinics scheduled, or marginally worse. Objective function two unsurprisingly performs the worst regarding number of clinics scheduled.

Figure 6.5: Number of Clinics Scheduled per Region per Objective Function

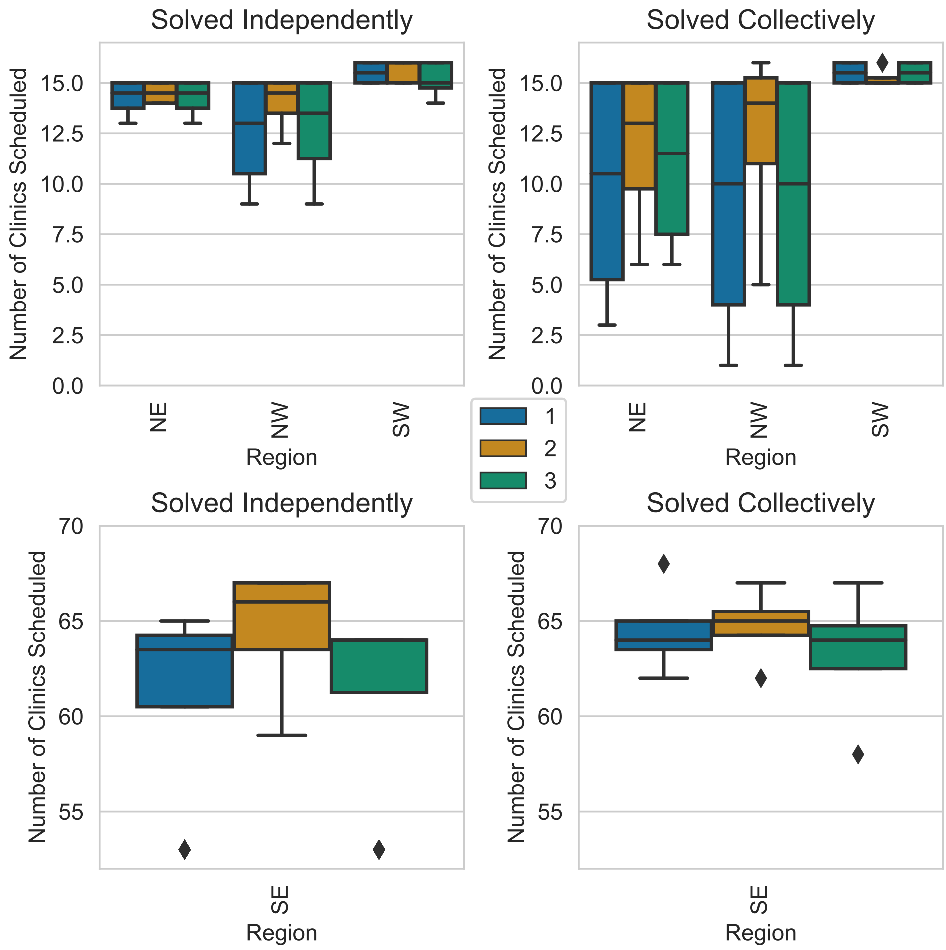


Table 6.9: Estimated Overcollection for All Wales - Actual WBS Figures Vs. Model

	Objective Function	Planning Horizon 1	Planning Horizon 2
Actual WBS Figures: Estimated Overcollection		1,080.0	63.0
Collective Model: Mean Estimated Overcollection	1	124	200
	2	0.0	0.0
	3	0.0	0.5
Independent Model: Mean Estimated Overcollection	1	388.0	371.0
	2	17.0	15.5
	3	18.0	15.5

Figure 6.6: Estimated Overcollection per Region per Objective Function (Solved Independently and All Wales Collectively)

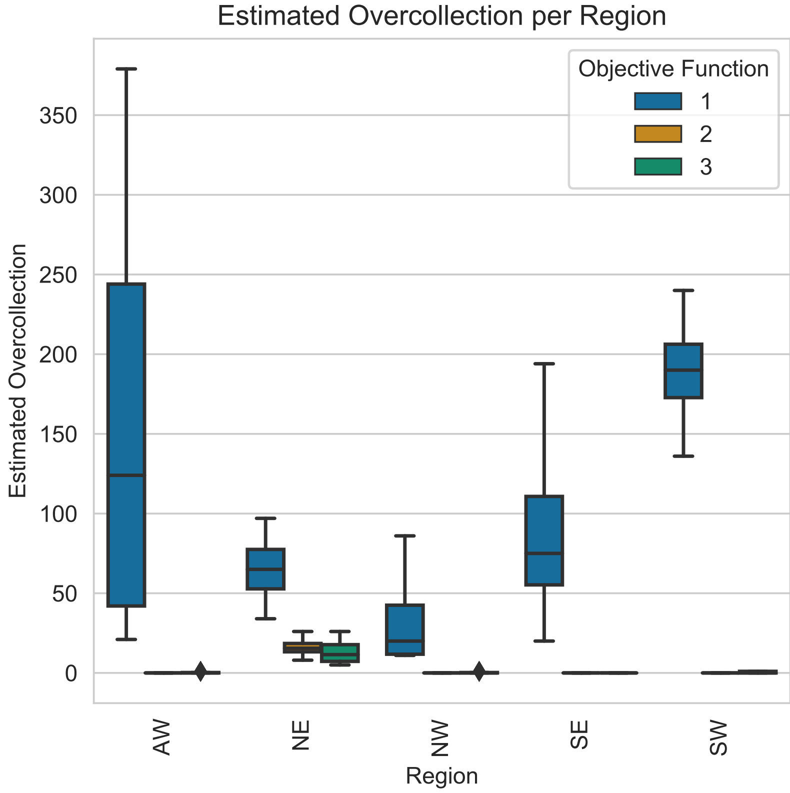


Figure 6.6 presents the estimated overcollection values for each region when solved independent of other regions and all of Wales when solved collectively, comparing solutions across alternative objective functions. It is clear from Figure 6.6 that objective function one performs poorly regarding minimising overcollection, and demonstrates that a focus on only minimising clinics often results in significant overcollection of blood donation units. Objective functions two and three perform similarly, with estimated overcollection being zero in most instances. The model fails to achieve a zero estimated overcollection figure for the North East region; this is caused by fewer available clinics in this region and a smaller average available time window size per clinic as displayed in Table 5.1, leaving a limited choice for the model.

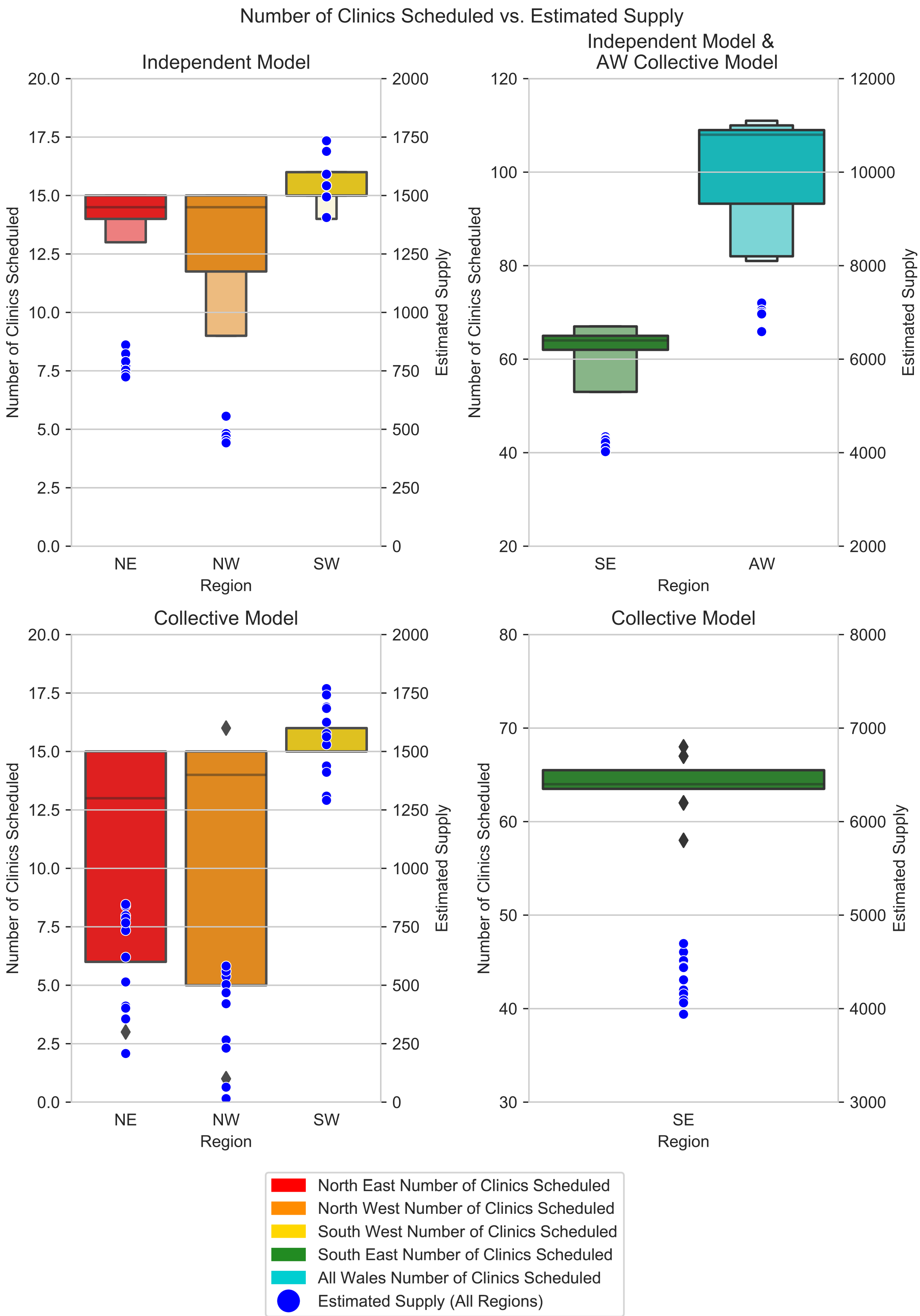
Due to demand and therefore overcollection being determined collectively across all regions in the AW instances, we cannot observe estimated overcollection per individual region in these cases.

Regions

Each of the four regions in Wales were tested for the same instances when solved individually and collectively. In the instances where regions are solved independently of each other, the demand for blood collection is predetermined per region according to the historical proportion of total blood collected that was from the corresponding region. Figure 6.7 illustrates the number of clinics scheduled for each region when solved independently (coloured boxplots) alongside the estimated supply from each region (blue plots). This figure shows that a similar number of clinics are required to be scheduled per four-week planning horizon for North East, North West, and South West regions to achieve the corresponding collection demand. However, the South West region provides significantly higher collection values than the two north regions. This relates to the efficiency of the clinics in different regions, as the South West region contains numerous clinics that collect high volumes of donations each time they operate. Due to the more rural nature of most clinics in both north regions, but particularly the North West region, this explains the low collection estimates from a similar number of clinics compared to the South West region.

Table 6.10 displays the mean number of clinics scheduled per region, when solved both independently and collectively, in addition to the mean estimated supply per region and the mean estimated supply per clinic per region. This effectively demonstrates the efficiency of collection in each region, with the North West having the lowest mean estimated supply per scheduled clinic with 35.7 units when solved independently of other regions, and 36.2 units when solved collectively. The South West region is by far the most efficient region for blood collection with 98.3 units estimated mean estimated supply per scheduled clinic when solved independently and 100.7 units when solved collectively. The mean estimated supply per scheduled

Figure 6.7: Number of Clinics Scheduled vs. Estimated Supply per Region (Solved Independently) and All Wales Collectively



clinic for the whole of Wales when all regions are solved collectively is 68.2 units, with a lower efficiency of 66 units per clinic when all regions are solved independently.

Table 6.10: Mean Scheduled Clinics and Supply per Region

	Region	Mean Number of Clinics Scheduled	Mean Total Supply	Mean Supply per Clinic
Solved Independently	NE	14.3	773.2	53.9
	NW	13.1	467.5	35.7
	SE	62.3	4,178.3	67.0
	SW	15.4	1,513.1	98.3
	AW	105.1	6,932.1	66.0
Solved Collectively	NE	10.8	606.8	56.0
	NW	10.1	364.8	36.2
	SE	64.2	4,325.2	67.4
	SW	15.4	1,552.3	100.7
	AW	100.5	6,849.1	68.2

Figure 6.8: Number of Clinics Scheduled - Actual Vs. Model

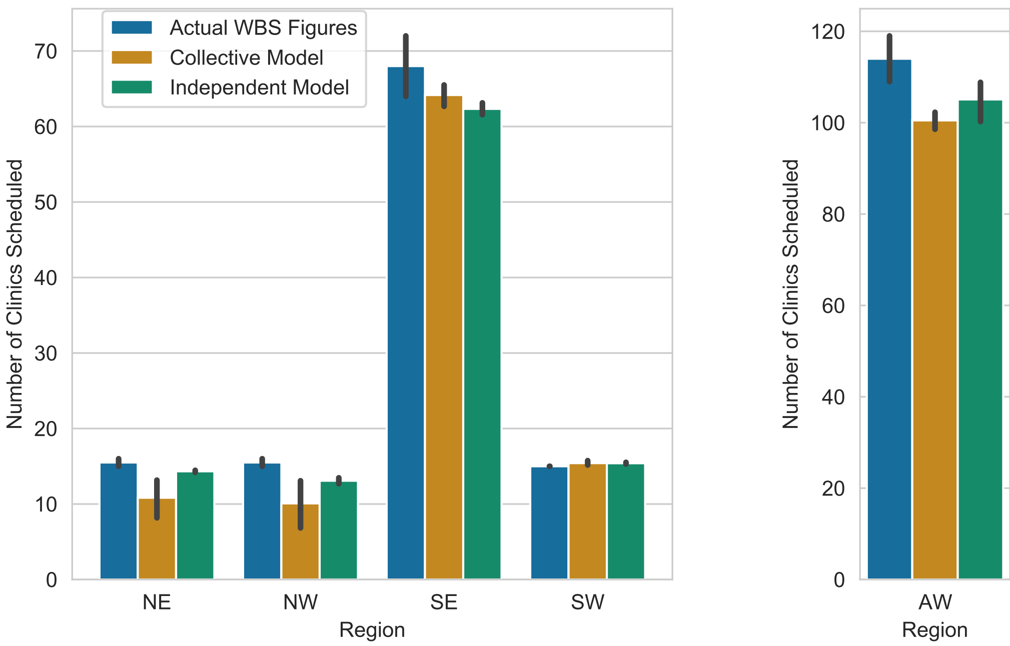


Table 6.11 presents the mean percentage of total blood collection provided by each region, comparing the figures from the WBS data for the two selected four-week planning horizons in 2019 with the mean percentages chosen by the BDCSM when

all regions are solved collectively. Both North East and North West regions are chosen to supply significantly less of the total collection by the collective model than the actual WBS mean percentages. This further supports the notion that the two north regions are inefficient.

Furthermore, comparing the actual number of scheduled clinics by the WBS during the same planning horizon periods as utilised by the BDCSM Model, as presented in Figure 6.8, it is evident that the model (both when all regions are solved collectively and independently) performs better than the manual scheduling process currently in place at the WBS. Overall, the collective model schedules the least clinics to satisfy the same demand as the independent model, and therefore would likely reduce clinic-associated costs. Figure 6.8 also illustrates that the collective model schedules fewer clinics for the two north regions, as it is more efficient to collect from the south regions. Conversely, the number of clinics scheduled for the South West region remains approximately the same for both models and for the actual WBS schedule. This highlights the effectiveness of collection in this region, with a minimum of 15 clinics scheduled in this region in both models where the maximum permitted due to constraints is 16 clinics.

Table 6.11: Percentage of Total Collection per Region - Actual WBS Figures Vs. Collective Model

		Actual	Collective Model		
Season	**Region**	**Percentage of Total Collection**	**Mean Percentage of Total Collection**	**Mean Percentage of Total Collection (Constraints Type 1)**	**Mean Percentage of Total Collection (Constraints Type 2)**
Planning Horizon 1 07/01/2019 - 03/02/2019	NE	11.4%	9.5%	11.7%	7.3%
	NW	8.6%	6.3%	7.9%	4.7%
	SE	59.1%	61.2%	58.7%	63.8%
	SW	20.9%	22.9%	21.6%	24.2%
Planning Horizon 2 22/07/2019 - 18/08/2019	NE	10.5%	8.1%	11.3%	4.9%
	NW	6.6%	4.2%	6.9%	1.6%
	SE	61.4%	65.2%	61.8%	68.6%
	SW	21.5%	22.5%	19.9%	25.0%

Table 6.12: Allocated Demand per Region for Region-Independent Model

	Allocated Demand in per Region			
	North East	**North West**	**South East**	**South West**
Planning Horizon 1 07/01/2019 - 03/02/2019	764	470	4,272	1,494
Planning Horizon 2 22/07/2019 - 18/08/2019	719	352	4,022	1,406

Seasonality

For the experimental results, two four-week planning horizons were considered: season one is a winter period (07/01/2019 – 03/02/2019) and a season two is a summer period (22/07/2019 – 18/08/2019). Demand varies slightly throughout a calendar year in addition to the availability of many clinics, with Table 6.12 displaying the allocated demand per planning horizon for the model. Observing the differences in solutions over these two planning horizons allows some insight into the effect of the season on collection.

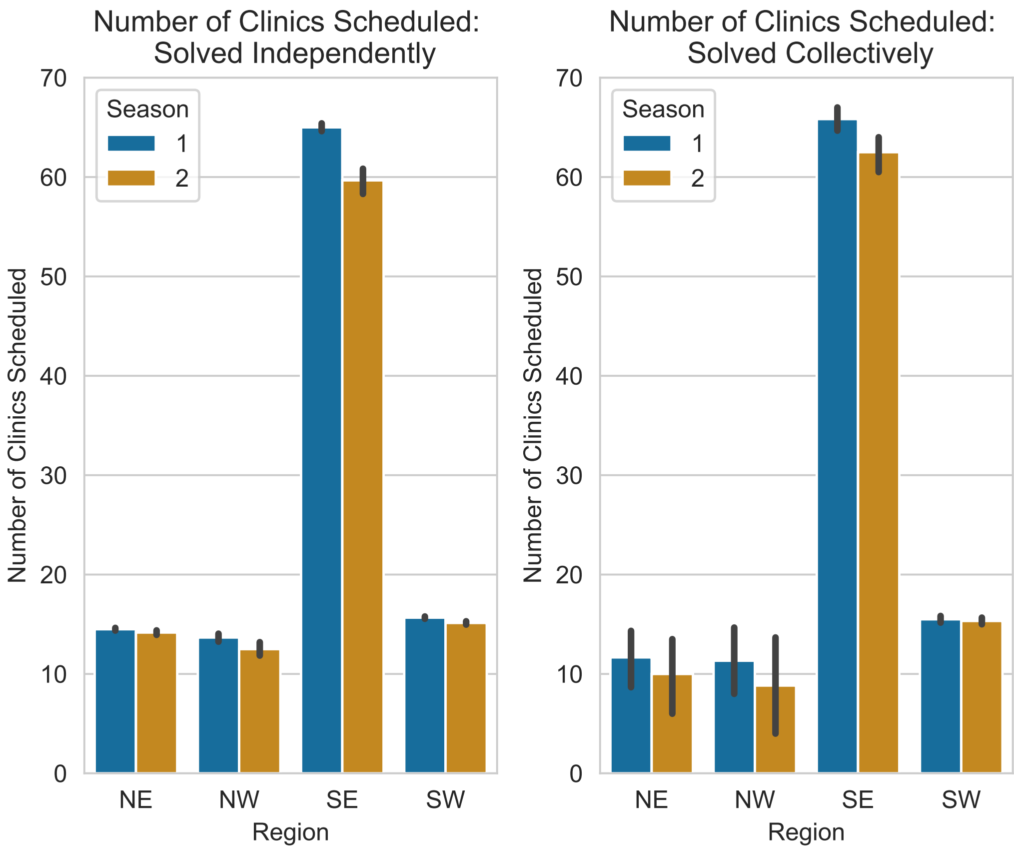
It is evident from Figure 6.9 that fewer clinics are scheduled for each region for season two than season one, and this is due to season two having a lower demand than season one and therefore less clinics are required to meet the demand. Whilst the number of clinics scheduled per region for season one is similar across both models with slightly fewer in the north regions, for season two the collective model chooses to schedule even fewer clinics in the north regions (shown by the 'error' bars to portray the variability across instances) and more clinics in the South East region than for season one. This demonstrates that the current practice of operating a consistent number of clinics per region for all planning horizons is an inefficient approach, and scheduling to demand could reduce costs.

Additional Constraints

Alternative scenarios are considered for the constraints of the model where a set of constraints are either included ('constraints type one') or excluded ('constraints type two') for different instances. The constraints involved are described in (3.17) and ensure that for each region, a given minimum number of clinics are scheduled per four-week planning horizon to provide adequate working hours for the clinic-based workers in each region.

It is clear from Figure 6.10 that constraints type two (the exclusion of the specific set of constraints) results in fewer clinics scheduled in both north regions, both in

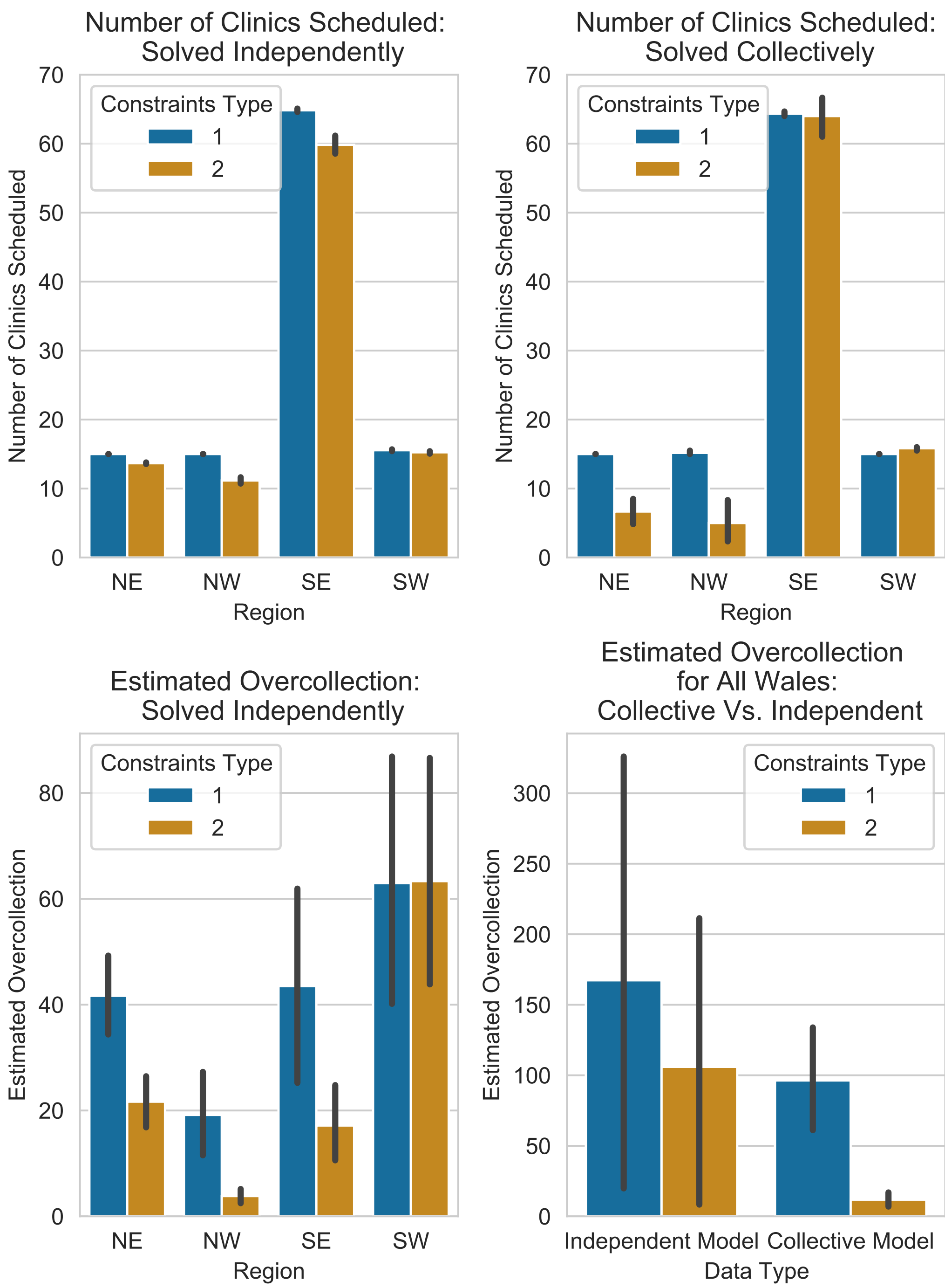
Figure 6.9: Number of Clinics Scheduled per Region per Season



the independent model and the collective model. This difference in the number of clinics scheduled for the two north regions is more significant in the collective model, as the demand is considered across the whole of Wales and thus the model prefers to collect from the south regions when it has the ability to do so. Regarding estimated overcollection, constraints type two performs better in most regions in the independent model, and significantly better overall for both models.

Table 6.11 details the percentage of total collection to be provided by each region for both the collective model and the actual historical figures for the WBS. These percentages are divided into the two planning horizons and then further into the two constraint types for each planning horizon. These figures give an insight into how the model chooses to distribute demand across the four regions compared to the WBS current practice. The model where regions are solved independently is not included in Table 6.11 as demand per region is pre-determined.

Figure 6.10: Number of Clinics Scheduled per Region per Constraints Type



The notable contrast between the two constraint types for the collective model are also evident in the percentage of total collection per region in Table 6.11. The mean percentage of estimated collection provided by each region for constraints type one is similar to the actual WBS figures, for both planning horizons. However, the corresponding percentages per region for constraints type two are considerably lower for both north regions, with a higher proportion of the total collection instead supplied by the two south regions.

6.2.3 Summary: Blood Donation Clinic Scheduling Model Results

This section has presented both the computational results and solutions to the Blood Donation Clinic Scheduling Problem from the model developed in Python, with regions solved both independently and collectively.

It is clear that objective function three is inferior in computational performance considering total run time compared to the other two alternative objective functions, across all regions. For objective functions one and two, all regions can still be solved (either collectively or independently) to optimality or near optimality (due to the optimality gap) in 90 minutes or less. However, since objective function three aligns both the WBS strategic aims of matching supply to demand by minimising overcollection, and provides more opportunity to reduce costs by minimising clinics, this objective function is preferable for the WBS.

The maximum total run time utilising objective function three for the whole of Wales can be as much as eight hours when all regions are solved collectively; this is still significantly more efficient than the current practice at the WBS where an original clinic schedule is manually created and takes an employee approximately five working days. This model provides an optimal (or near optimal) clinic schedule which can be altered accordingly depending on any changes in circumstances or venue availability. This is one way in which we have answered research question one: 'How can mathematical modelling help to schedule the Welsh Blood Service's blood donation clinics more efficiently?'

Generally, utilising objective function three for both the independent and collective models improves upon the actual Welsh Blood Service figures for overcollection (shown in Table 6.9) and number of clinics scheduled (shown in Figure 6.8) for the same planning horizons. However, the demand determined for the model is based on the analysis of the issuing data in the weeks subsequent to the selected planning horizons; this is not known at the time of creating clinic schedules at the Welsh Blood Service and thus, the comparison of overcollection between actual WBS figures and the model is not reliable. Despite this, the model still schedules fewer clinics (as much as up to 20 fewer clinics when regions are solved collectively) to collect a similar amount of blood as the actual WBS collection over the same four-week periods. This demonstrates that the efficiency of clinic schedules can be improved significantly by utilising mathematical modelling, in response to research question one.

Additionally, some insight is gained into further potential improvement in efficiency in blood collection by the WBS regarding the restriction of scheduling a minimum number of clinics in each region for each planning horizon. Understandably, the WBS include this constraint when scheduling clinics to provide consistent work for clinic-based workers in all regions. However, from these experimental results it is evident that this collection model results in an increase in the number of clinics scheduled due to the low-yield of clinics in the north regions, particularly the North West region. The effect on workers of removing this constraint is further discussed in the experimental results for the Blood Collection Workforce Scheduling Model, which will be presented next.

6.3 Blood Collection Workforce Scheduling Model

Stage two of our mathematical model is the Blood Collection Workforce Scheduling Model (BCWSM) which was also developed in Python, with the development detailed in Chapter 5. This model was run for each test instance described in Table 5.2 for each region (both when solved collectively with other regions in stage one i.e. 'AW' instances, and when solved independently) for each of the two alternative objective functions. The model utilises an optimality gap of 1% i.e. a solution may be considered optimal if it is within 1% of the best known solution; this is to enable a solution to be reached within a smaller time and to avoid situations where the processor may run out of working memory. A larger optimality gap was considered for stage two of the model, compared to stage one, to cope with the increase in complexity. From experimental runs of the BCWSM using an optimality gap of 0.5%, for larger instances such as the South East region, the solver occasionally ran out of working memory. Increasing the optimality gap from 0.5% to 1% solved this issue. Each instance was run for numerous iterations with the removal of one worker after each iteration, until the problem became infeasible to solve. The experimental results are presented in this section, along with explanation of how the model can increase efficiency of the workforce scheduling process for the WBS, in answer to research question two (described in Section 1.5).

From initial experimental results of the original model formulation described in Section 4.3, it became evident that the workforce schedule solutions of the model were not realistic. Many workers were scheduled their full contracted hours over the duration of the four-week planning horizon, with some weeks significantly in excess of their weekly contracted hours, counteracted by working fewer hours in other weeks. These hours in excess of a worker's weekly contracted hours are not considered overtime by the WBS as overtime is calculated over the four-week planning horizon. Additionally, some workers were consistently scheduled either no or very few hours each week of the planning horizon. The Welsh Blood Service workforce planning team attempt to assign workers their full contracted hours (including any annual leave and training) each planning horizon, to maintain employee satisfaction and to reduce wastage of monetary costs as they are required to pay the workforce for their contracted hours whether they are worked or not. The original model formulation did not maximise the utilisation of contracted hours nor did it maximise fairness in distribution of working hours across the workforce. To enable the model to consider these two factors, the modifications described in Section 4.4 were made to the

BCWSM, and the improvement in results are displayed in Appendix C alongside the difference in computation times. Therefore, the results presented in this section are that of the modified version of the BCWSM.

6.3.1 Computational Results

The computational results of all test instances are now discussed to observe if the model provides a solution in a feasible timeframe for potential implementation at the Welsh Blood Service. Tables 6.13– 6.16 detail the computational results for each instance under objective function one, including the minimum and maximum solution times of all iterations per instance. The minimum and maximum number of decision variables and constraints per instance are also displayed in these tables to illustrate the size of each problem. As a reminder to the reader, the instance nomenclature for the BCWSM follows the structure of A/B/C/D/E/F/G/H where:

- A denotes the region considered i.e. {NE = North East; NW = North West; SE = South East; SW = South West} where 'AW' prior to this indicates instances where the region was solved collectively with other regions in the BDCSM (stage one)
- B denotes the planning horizon i.e. {1 = winter; 2 = summer}
- C denotes the inclusion of the additional constraints in the BDCSM i.e. {1 = minimum number of clinics to be scheduled per region included; 2 = minimum number of clinics to be scheduled per region excluded}
- D denotes the objective function of the BDCSM i.e. {1 = minimisation of clinic days scheduled; 2 = minimisation of estimated overcollection; 3 = minimisation of both clinic days scheduled and estimated overcollection}
- E denotes the objective function of the BCWSM i.e. {1 = minimisation of total cost of overtime; 2 = minimisation of total cost of scheduled hours}
- F denotes the inclusion of driving role constraints and modes i.e. {1 = included;

2 = excluded}

- G denotes the inclusion of the prioritisation of dummy workers removed before actual WBS workers i.e. {1 = included; 2 = excluded}
- H denotes the consideration of sick leave i.e. {1 = not considered; 2 = considered}

The North East, North West and South West regions have significantly fewer decision variables and constraints than the South East region, and thus take significantly less time to run; the total run time for any of these smaller regions is less than 600 seconds which means that the model can produce a workforce schedule for each of these three regions in fewer than 10 minutes. Due to the size of the problems for the South East region, the maximum total run time is just over 43,000 seconds ($\approx$ 12 hours). However, as displayed in Figure 6.11, the total run time is less than 10,000 seconds (less than three hours) at the current workforce level (at the 77th iteration) and remains to be the case for some iterations thereafter. Therefore, the BCWSM requires considerably less time than it takes the WBS to create the clinic workforce schedules currently, with the completion of four weekly rotas requiring approximately 32 hours for both the South East and South West regions collectively, and approximately 16 hours for both the North East and North West regions collectively. However, these WBS figures include the time required to input the annual leave of workers which would also need to be considered in addition to the run times of the model.

The results included in Tables 6.13– 6.16 are from the original test instances where a worker is removed at random at each iteration. To enable the observation of the direct effect of various parameters on the goodness of solutions across test instances, we run the models again using an identical workforce across all instances, with one worker still removed at each iteration, but this worker is the same for all instances. In practice, the workforce included at each iteration of the 1131111 instance in the original run of the modified model is input into each corresponding iteration of all other instances for a given region. The instances where there is no prioritisation of the removal of dummy workers over actual WBS workers is not included as this variable method can not be put utilised whilst inputting a predetermined workforce

Table 6.13: North East Region Computational Results

Instance	Total Number of Iterations	Number of Decision Variables		Number of Constraints		Solution Time (secs)		Total Run Time (secs)	
		Min.	**Max.**	**Min.**	**Max.**	**Min.**	**Max.**	**Min.**	**Max.**
NE1131111	19	47,160	106,896	29,291	62,678	10.4	25.1	189.4	408.3
NE1131112	19	47,160	106,896	29,291	62,678	7.2	18.8	148.6	335.2
NE1131121	18	50,304	106,896	31,048	62,678	7.2	17.8	155.2	330.6
NE1131211	19	23,640	53,584	15,012	32,438	4.7	11.9	45.7	97.2
NE1132111	19	47,160	106,896	29,293	62,678	9.2	23.2	191.2	401.9
NE1231111	19	45,480	103,088	27,417	58,674	7.7	21.1	169.8	384.3
NE1232111	19	45,480	103,088	27,416	58,674	8.1	21.5	165.6	376.4
NE2131111	20	44,016	106,896	27,536	62,675	10.0	26.9	234.6	557.3
NE2132111	19	47,160	106,896	29,290	62,675	12.3	28.7	281.7	587.6
NE2231111	21	37,960	99,280	22,473	54,667	8.1	25.7	153.6	512.3
NE2232111	20	40,880	99,280	24,007	54,667	7.8	26.7	153.7	531.4
AWNE1131111	17	53,448	106,896	32,314	61,692	11.9	38.9	163.4	333.4
AWNE1131112	17	53,448	106,896	32,314	61,692	7.9	21.7	138.8	279.6
AWNE1131121	16	56,592	106,896	34,042	61,692	9.4	22.9	142.3	267.6
AWNE1131211	17	26,792	53,584	16,354	31,452	5.3	13.6	40.8	83.2
AWNE1132111	17	53,448	106,896	32,314	61,692	11.6	31.8	157.2	320.7
AWNE1231111	21	30,680	80,240	13,836	33,664	6.4	18.5	74.6	178.1
AWNE1232111	21	30,680	80,240	13,837	33,664	5.9	17.5	98.3	183.9
AWNE2131111	17	53,448	106,896	32,314	61,689	14.4	34.5	190.7	421.8
AWNE2132111	19	47,160	106,896	28,859	61,689	10.7	30.3	169.1	391.4
AWNE2231111	20	29,904	72,624	11,252	25,653	4.5	12.9	76.7	187.7
AWNE2232111	20	29,904	72,624	11,253	25,653	4.3	12.6	76.9	172.0

Table 6.14: North West Region Computational Results

Instance	Total Number of Iterations	Number of Decision Variables Min.	Number of Decision Variables Max.	Number of Constraints Min.	Number of Constraints Max.	Solution Time (secs) Min.	Solution Time (secs) Max.	Total Run Time (secs) Min.	Total Run Time (secs) Max.
NW1131111	13	40,872	81,744	25,775	48,617	7.1	17.1	160.0	313.7
NW1131112	11	47,160	81,744	29,291	48,617	6.3	12.2	142.2	244.7
NW1131121	11	47,160	81,744	29,291	48,617	6.8	14.7	149.9	255.6
NW1131211	12	22,064	40,976	14,094	25,097	2.9	7.3	39.3	71.8
NW1132111	13	40,872	81,744	25,773	48,617	7.2	17.8	158.7	299.2
NW1231111	13	36,504	73,008	20,819	39,293	3.3	13.9	117.7	238.9
NW1232111	13	36,504	73,008	20,824	39,293	6.1	14.0	131.7	242.7
NW2131111	13	40,872	81,744	25,779	48,610	9.0	17.4	163.5	300.7
NW2132111	13	40,872	81,744	25,781	48,610	8.2	20.4	158.1	300.3
NW2231111	14	29,664	64,272	14,780	29,962	4.3	10.2	104.1	206.0
NW2232111	14	29,664	64,272	14,780	29,962	5.1	11.2	101.5	197.5
AWNW1131111	12	44,016	81,744	27,130	47,863	11.2	20.5	162.5	277.0
AWNW1131112	11	47,160	81,744	28,862	47,863	8.3	15.6	122.1	201.9
AWNW1131121	12	44,016	81,744	27,133	47,863	7.4	14.3	112.0	204.7
AWNW1131211	12	22,064	40,976	13,688	24,343	4.5	8.6	40.6	60.7
AWNW1132111	12	44,016	81,744	27,127	47,863	11.6	19.9	176.6	304.0
AWNW1231111	11	30,360	52,624	10,096	16,783	5.6	10.1	73.6	118.5
AWNW1232111	10	32,384	52,624	10,705	16,783	7.1	12.0	87.0	141.6
AWNW2131111	12	44,016	81,744	27,129	47,856	11.1	24.8	180.6	323.1
AWNW2132111	13	40,872	81,744	25,400	47,856	10.2	23.3	143.1	297.8
AWNW2231111	15	17,336	40,976	1,949	4,344	2.0	5.8	35.9	78.9
AWNW2232111	15	17,336	40,976	1,948	4,344	2.0	6.1	35.5	76.5

Table 6.15: South East Region Computational Results

	Number of Iterations	Number of Decision Variables		Number of Constraints		Solution Time (secs)		Total Run Time (secs)	
Instance		**Min.**	**Max.**	**Min.**	**Max.**	**Min.**	**Max.**	**Min.**	**Max.**
SE1131111	99	400,840	1,122,352	336,427	923,630	310.3	36,320.4	7,233.9	41,409.7
SE1131112	98	408,128	1,122,352	342,353	923,625	191.3	22,547.0	5,085.8	26,654.2
SE1131121	102	378,976	1,122,352	318,631	923,630	215.9	36,276.4	4,968.4	40,260.8
SE1131211	100	196,992	561,792	166,665	467,202	81.7	36,327.5	962.2	37,614.4
SE1132111	99	400,840	1,122,352	336,423	923,630	312.3	36,426.5	9,114.1	43,409.6
SE1231111	98	414,400	1,139,600	348,814	941,074	230.3	36,340.0	6,588.8	43,231.4
SE1232111	90	473,600	1,139,600	397,162	941,074	459.3	36,280.3	6,714.7	41,718.0
SE2131111	91	437,976	1,070,608	362,054	871,138	277.3	36,229.3	5,710.3	40,561.9
SE2132111	90	444,928	1,070,608	367,641	871,138	222.5	13,426.6	5,386.5	18,974.2
SE2231111	91	367,416	898,128	289,532	696,700	111.0	3,028.9	3,073.2	11,052.8
SE2232111	98	326,592	898,128	258,200	696,700	105.4	1,455.0	2,471.4	11,647.8
AWSE1131111	97	383,496	1,036,112	313,732	831,944	190.7	36,613.5	5,904.7	41,576.1
AWSE1131112	99	370,040	1,036,112	303,039	831,941	160.5	36,262.0	4,078.2	39,587.0
AWSE1131121	97	383,496	1,036,112	313,736	831,944	200.9	5,528.9	4,140.9	12,548.3
AWSE1131211	96	195,344	518,672	160,330	419,196	78.9	36,284.0	761.2	37,177.5
AWSE1132111	99	370,040	1,036,112	303,057	831,944	254.1	9,471.0	6,499.4	15,563.8
AWSE1231111	90	487,936	1,174,096	410,038	971,496	268.9	364,54.2	6,657.9	42,167.8
AWSE1232111	91	480,312	1,174,096	403,800	971,496	206.1	36,366.2	6,780.7	41,785.4
AWSE2131111	92	389,360	967,120	311,075	760,156	196.2	36,497.7	4,602.8	41,587.2
AWSE2132111	95	370,520	967,120	296,424	760,156	169.4	15,326.2	3,607.0	18,590.8
AWSE2231111	102	320,736	949,872	256,917	744,720	126.4	15,716.0	3,013.8	18,166.5
AWSE2232111	91	388,584	949,872	309,528	744,720	154.9	7,877.4	3,747.8	11,667.3

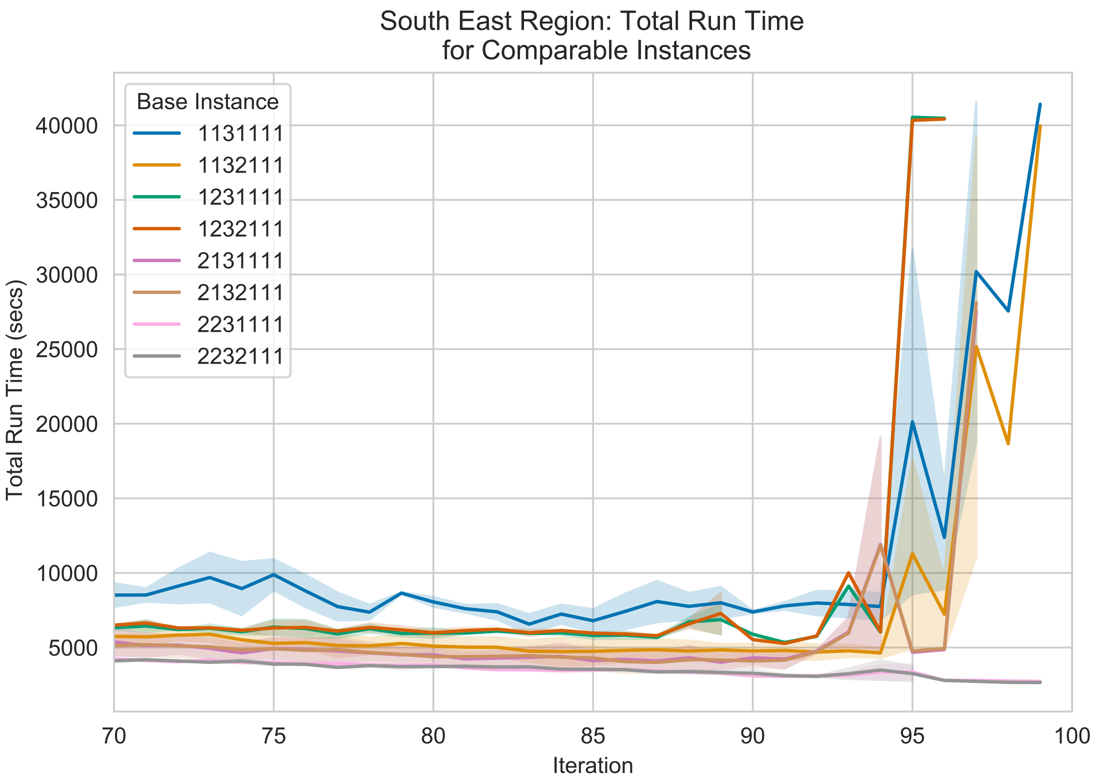
Table 6.16: South West Region Computational Results

	Number of Iterations	**Number of Decision Variables**		**Number of Constraints**		**Solution Time (secs)**		**Total Run Time (secs)**	
Instance		**Min.**	**Max.**	**Min.**	**Max.**	**Min.**	**Max.**	**Min.**	**Max.**
SW1131111	19	40,120	84,960	25,707	51,154	12.3	19.1	133.0	255.8
SW1131112	16	47,200	84,960	29,726	51,153	8.0	15.0	126.8	231.5
SW1131121	19	40,120	84,960	25,702	51,154	7.7	15.2	104.3	219.1
SW1131211	19	26,792	56,736	16,887	34,354	5.6	16.8	58.5	113.9
SW1132111	19	40,120	84,960	25,709	51,154	5.4	19.8	151.7	360.1
SW1231111	18	43,992	87,984	28,755	54,374	14.2	28.6	195.9	425.0
SW1232111	18	43,992	87,984	28,755	54,374	11.1	23.9	141.4	380.1
SW2131111	18	42,480	84,960	27,047	51,150	9.4	18.1	130.4	258.1
SW2132111	18	42,480	84,960	27,047	51,150	9.2	17.4	129.9	258.1
SW2231111	18	40,968	81,936	25,339	47,930	8.9	20.6	132.1	247.3
SW2232111	18	40,968	81,936	25,339	47,930	9.3	20.3	133.3	249.7
AWSW1131111*	18	39,456	78,912	23,175	43,802	7.3	15.4	78.8	149.2
AWSW1131112*	17	41,648	78,912	24,321	43,801	7.4	14.4	80.6	149.1
AWSW1131121*	19	37,264	78,912	22,030	43,802	8.2	15.5	79.3	151.2
AWSW1131211*	18	26,352	52,704	15,132	29,170	5.2	11.5	44.0	83.0
AWSW1132111*	18	39,456	78,912	23,176	43,802	7.9	13.0	80.4	150.5
AWSW1231111*	12	56,640	84,960	34,475	50,242	8.5	15.3	117.6	169.3
AWSW1232111*	12	56,640	84,960	34,476	50,242	9.3	15.0	116.4	169.3
AWSW2131111	18	42,480	84,960	26,489	50,034	11.5	25.7	114.9	218.2
AWSW2132111	18	42,480	84,960	26,489	50,034	11.6	24.7	115.0	216.2
AWSW2231111	17	46,436	87,984	29,647	53,362	11.4	422.7	127.6	530.0
AWSW2232111	17	46,436	87,984	29,648	53,362	10.6	50.9	125.5	233.9

* instances with annual leave constraints removed.

at each iteration. When data from these comparable results are being referred to, it will be explicitly mentioned such as results displayed in Figure 6.11.

Figure 6.11: Total Run Times for South East Region Comparable Instances



Instances for the South Wales region where it is solved collectively with other regions in stage one of the model (i.e. 'AWSW' instances) during the winter planning horizon (season one) are infeasible from the very first iteration, where there are duplicates of every worker included. This was found to be caused by the randomly generated annual leave schedules clashing with the scheduling of a clinic tour, as each clinic tour requires the same workers to be assigned to each day of the tour. The predetermined annual leave of the Registered Nurses prevented this constraint from being met. In practice, the Welsh Blood Service would likely not approve annual leave that would cause major disruption to the operation of a clinic that is particularly difficult to schedule, such as tour clinics. Alternatively, perhaps adjustments would be made to the number of chairs operated during the clinics to decrease the number of workers required: e.g. by reducing a tour clinic to nine or fewer chairs, only one nurse would be required to be assigned to the clinic tour. Alternatively, a nurse from another region (usually the South East due to geographical proximity) would be assigned to clinics in the South West region to assist with short-staffing situations. It is complex to generate annual leave schedules that are realistic for the WBS, and we were unable to receive their data for this due to complications with anonymity of workers and the data from the time of the considered planning horizons is not digitally stored and therefore would be difficult and time consuming to locate.

For the purpose of observation, we remove the annual leave constraint from the model and run these instances again, and denote these instances with an asterisk. As presented in Table 6.16, many of these instances fail to reach the 18th iteration where the current workforce level is considered; these cases are mostly instances where the additional constraints ensuring a minimum number of clinics are scheduled per region in stage one (the BDCSM) are excluded, allowing the BDCSM to schedule a greater number of clinics and/or larger clinics in the South West region instead of the less efficient north regions. The remaining case where the 'AWSW' instance fails to reach the current workforce level is when sickness leave is included. These results infer that the current workforce in the South West region (at the time that the workforce data was shared) are inadequate to cope with both sickness leave and a maximisation of clinic scheduling efficiency in terms of collecting a greater proportion of blood donations from this region.

Objective Function

Recall in Section 4.4.2 we introduce the following two alternative objective functions for the modified version of the BCWSM; the minimisation of the total cost of overtime (4.34), and the minimisation of the total cost of scheduled hours (4.35). Both of these objective functions also minimise weekly overtime and undertime per clinic-based worker, and the penalty function to discourage scheduling a Deputy Supervisor to the clinic role of Supervisor if there is a Supervisor available.

Figure 6.13 contains all computational results from the modified model in the comparable instances, excluding instances that occurred for objective function one but not objective function two i.e. the instances with alternative parameters at stage two; This is to allow a fair comparison between the performance of the two objective functions. This figure illustrates the solution and total run times per region per objective function considered (at stage two of the model). Generally, there is no noticeable difference in computational performance between the two objective functions, despite objective function two minimising the total cost of scheduled hours in addition to scheduled overtime hours.

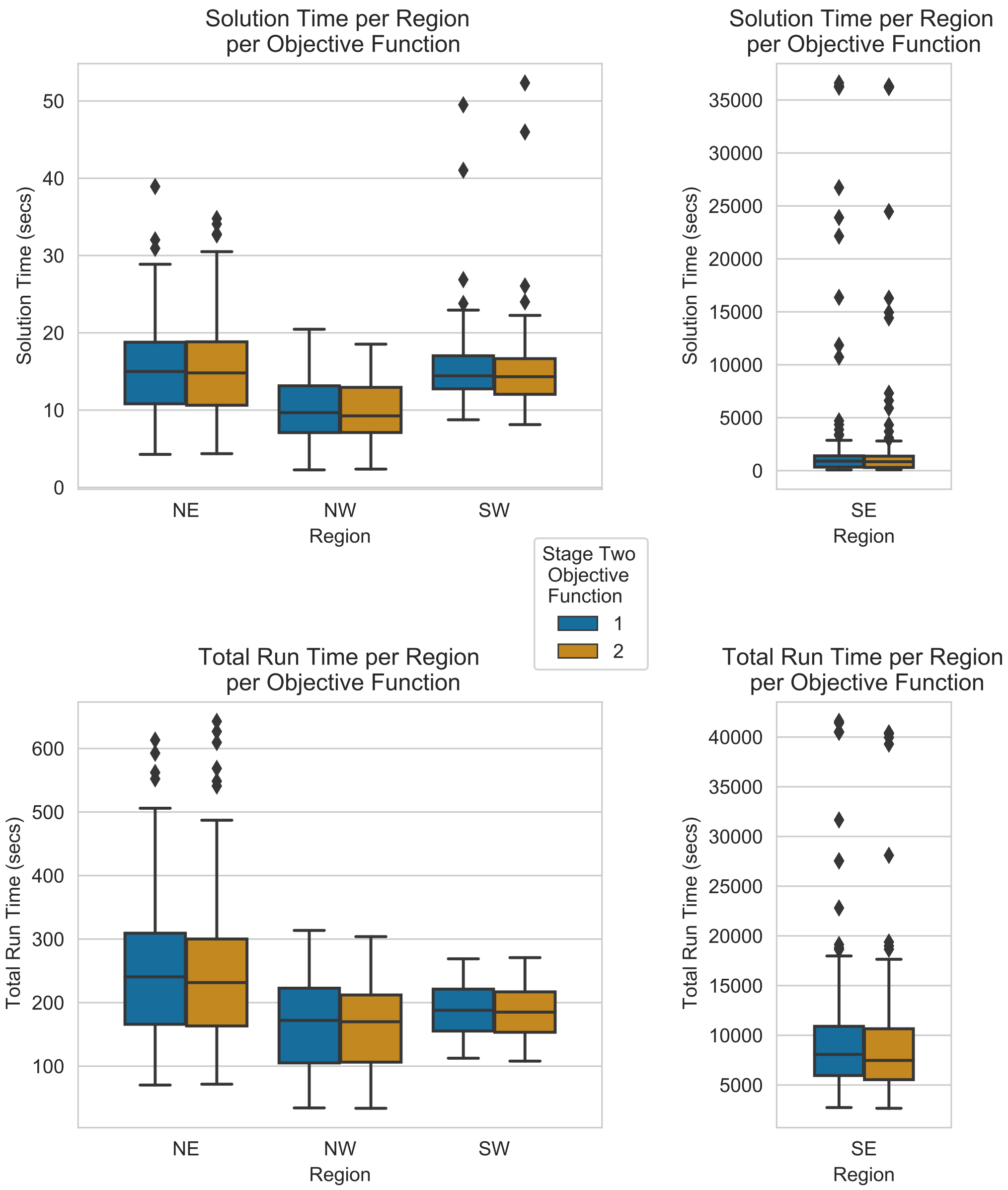
Alternative Parameters

In the BCWSM, three variable parameters are included in addition to the alternate objective functions; these will be described as 'driver constraints', 'dummy removal prioritisation' and 'sickness consideration', with each discussed in more detail in Section 5.7.3. The only parameter of these that causes the most noticeable difference in the computational time is the driver constraints. Due to the limitations of formulating the BCWSP in Python and PuLP, it is required that each worker has the same number of modes to enable the model to correctly iterate over the decision variables. This results in the model having to loop over each constraint $max\{\mathcal{M}\}$ times for each worker, where $max\{\mathcal{M}\}$ denotes the maximum number of modes of all workers in a given instance. Some workers may have up to five working modes if they are a driver of several vehicle types, and this can increase to 10 modes if they are a Deputy Supervisor due to the option of two clinic roles. This can cause a significant increase in the total run time of the model, especially for instances with a large number of workers, such as those considering the South East region.

To observe the effect on the model when there are no driver-related constraints, the resource demand constraints for drivers are removed in instances with driver constraints type two in addition to all driving roles removed from each worker's mode matrix. This reduces the maximum number of modes to be at most two –

this is the case for Deputy Supervisors with two possible clinic roles, with all other workers having at most one mode.

Figure 6.12: Computational Times per Region per Objective Function



Note: Only instances that were run for both objective functions are included i.e. instances with alternative parameters in stage two are excluded.

Figure 6.13: South East - Computational Times of Comparable Instances

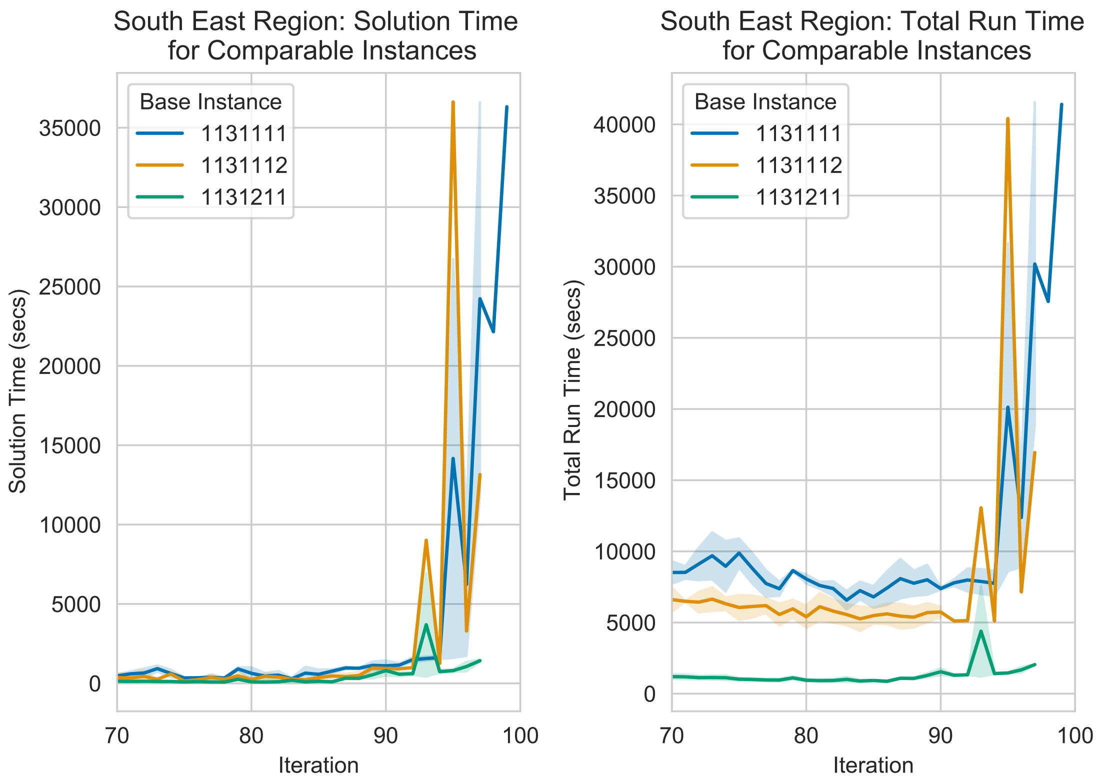


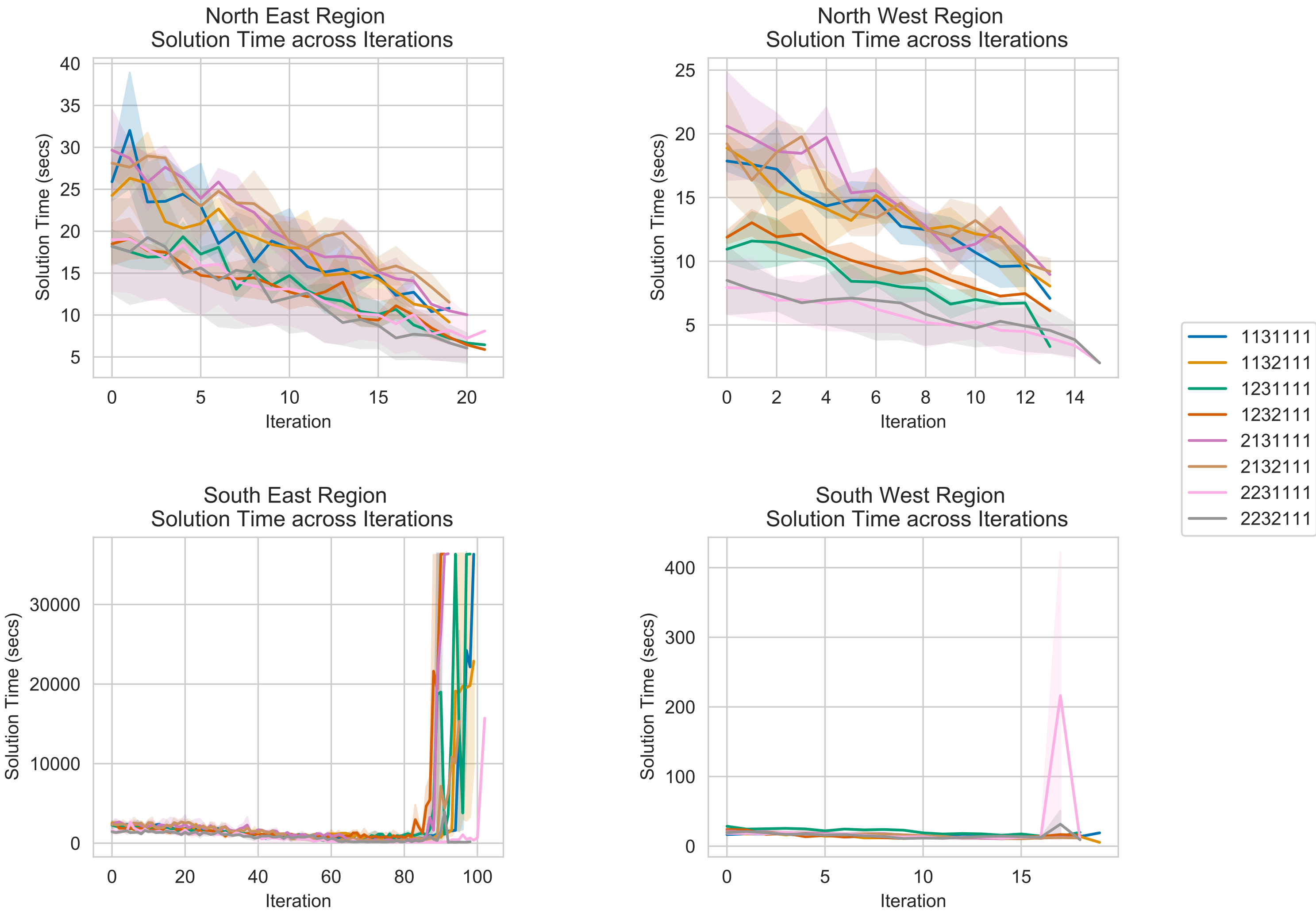
Figure 6.13 presents the solution and total run times of the comparable instances that include these alternative parameters, namely sickness consideration (1131112) and driver constraints (1131211), alongside the standard version of these instances (1131111) for the South East region. Only iterations from 70 onwards are displayed since the problem becomes increasingly more difficult as more workers are removed, with results prior to this following the same pattern as between iteration 70 to 90. Iteration 77 represents the current workforce level at the WBS in this region. It is evident from Figure 6.13 that the exclusion of driver-related constraints and driving role modes from the model drastically reduces the total run time of an instance across all iterations. Since the time taken to solve the problem is not notably reduced across all iterations when driver constraints are excluded, this infers that the problem does not become notably easier to solve in most cases, but does require significantly less time to 'set-up' the problem due to approximately half the number of constraints and decision variables, as shown in Tables 6.13– 6.16.

Complexity

For each instance, the number of decision variables and constraints changes for each iteration due to the removal of a worker at random, with the first iteration of each instance having the maximum figures for decision variables and constraints displayed in Tables 6.14– 6.16, and the final iteration of each instance having the minimum values for these figures. However, these values are not a good indication of the complexity of a given problem, and instead, the time taken to reach an optimal solution provides more insight into the complexity.

Figure 6.14 illustrates the time taken for the model to reach an optimal (or near optimal given the optimality gap) solution across iterations of each instance per region, where the iteration number is directly related to the size of the workforce considered; as the iteration number increases, the size of the workforce decreases. Both of the north regions display a trend of decreasing solution time as the number of workers decreases, however the solution time still remains less than a minute for the slower solution times. This trend is likely due to at most one clinic being scheduled per day of a planning horizon in both the North East and North West regions and at the first iteration, the larger workforce provides more potential solutions for the model to navigate. The South West region follows a similar trend with the exception of the 2231111 instance, where there is a significant increase in solution time at the 17th iteration caused specifically by the AWSW2231111 instance, after which the 'AW' version of the problem instance becomes infeasible, and the SW2231111 solution time remains small across all iterations. This increase in the case of the 'AW' version of the instance is likely caused by the greater number of CCA/CSA shifts required by the clinic schedule – a total of 155 shifts compared to 137 shifts for the SW2231111 instance.

Figure 6.14: Solution Time per Region across Iterations



Contrastingly, the solution time across iterations for South East region instances remains relatively low until later iterations, where the solution time increases drastically to over 30,000 seconds before the problem soon becomes infeasible. This demonstrates the increasing complexity of the problem once the workforce has decreased past the current workforce level at the WBS (which is reached at the 77th iteration). The sudden increase in solution time varies between instances, depending on a combination of both the details of the instance (such as clinic schedule) and the specific workforce considered at each iteration differing across instances.

6.3.2 Schedule Solutions

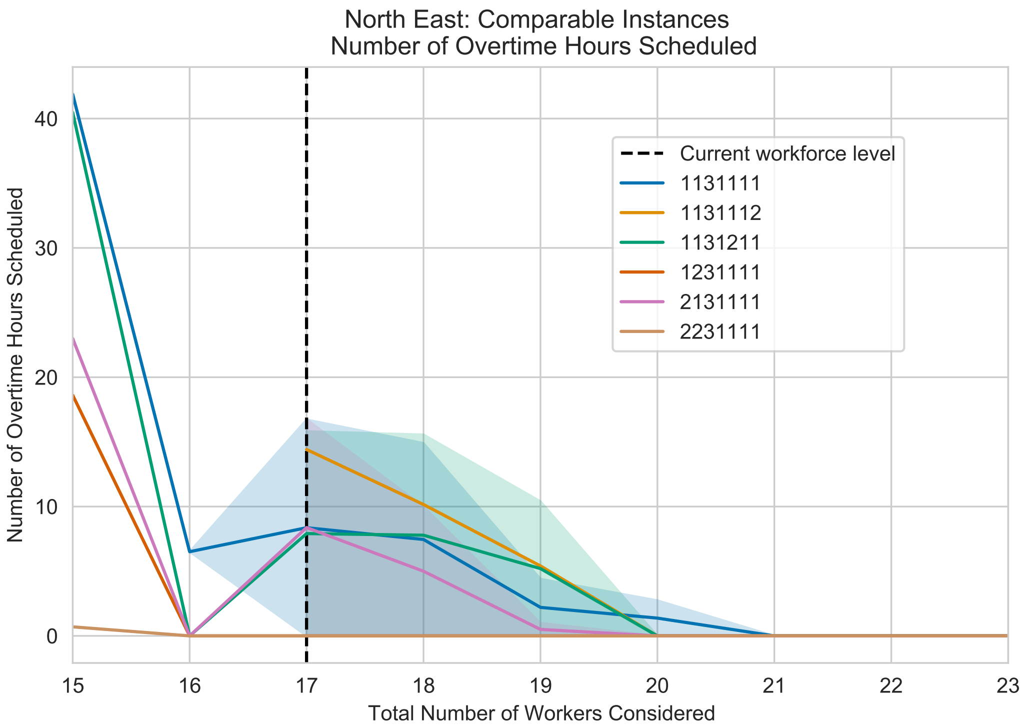
The solutions of the modified Blood Collection Workforce Scheduling Model are now discussed to observe how well the model performs at providing a feasible and realistic workforce schedule for the Welsh Blood Service, in addition to minimising costs according to the objective function considered.

Overtime

It is a main goal of the Welsh Blood Service to reduce paid overtime and maximise the usage of contracted hours, hence the decision to include an objective function that seeks to minimise the total cost of overtime (4.34). To observe the relationship between the number of workers considered in a test instance and the total number of overtime hours scheduled over the whole planning horizon, we present Figures 6.15–6.18 and discuss each corresponding region consecutively.

North East: For the North East region instances where stage one of the model is solved collectively with other regions, they become infeasible near the current WBS workforce level. This can be seen in Figure 6.15 where the number of overtime scheduled increases as workers are removed, for the majority of instances, until the current workforce level is reached and the problem becomes infeasible. This is likely due to how the generated annual leave and training schedules interact with the scheduled clinics, resulting in an inadequate number of workers available to cover absences. After the current workforce line, only the independent model data instances remain and the number of overtime scheduled decreases in the next iteration with the exception of the 1131112 instance (where sickness is included); this becomes infeasible at a workforce size of 16 workers which implies that the current workforce level may be too low to cope with periods of an above average sickness rate.

Figure 6.15: North East Region - Total Overtime Hours Scheduled for Comparable Instances



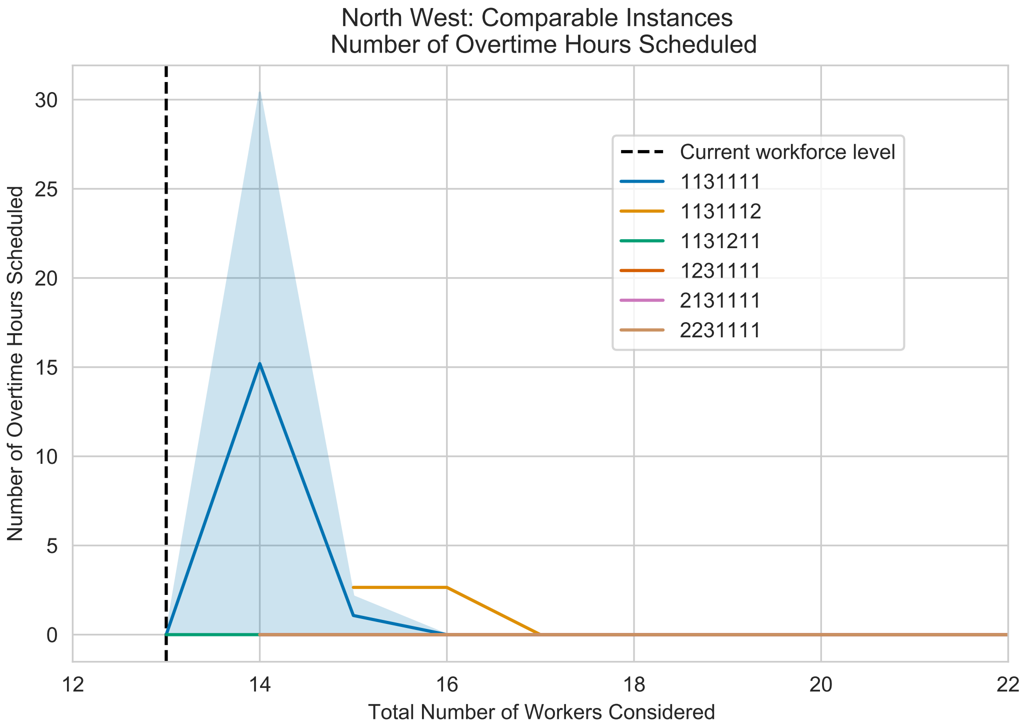
Please note: instance 1231111 remains at zero overtime until the workforce reduces to 15 workers.

With the exception of instances 1231111 and 2231111, overtime is required for a workforce smaller than 20 workers to enable the operation of all scheduled clinics at the standard donor capacity. This suggests that the current workforce level at the WBS is too low for the North East region if they are to continue with their pre-COVID-19 collection strategy. However, overtime remains at zero at the current workforce level for the two instances where the constraint for the minimum number of clinics to be scheduled per region is excluded in the stage one model; this is due to fewer clinics being scheduled over the four-week planning horizon in the North East region. If the WBS were to adjust their collection strategy to reduce the number of clinics required to be operated in the North East region each planning horizon, then the current workforce level would be sufficient and could possibly be reduced.

North West: For the North West region, Figure 6.16 depicts the lack of necessary overtime to ensure that all scheduled clinics in the region are adequately staffed. Overtime is only scheduled for instances 1131111 and 1131112. The latter of these two instances is the case where sickness leave is included, and this instance becomes infeasible for a workforce smaller than 15 workers which implies that the current workforce level of 13 clinic-based workers is not sufficient for the North West region to cope with sick leave without reducing the donor capacity of clinics and therefore collecting less blood than required to meet demand. This is reinforced by the fact that no instances remain feasible once the workforce decreases by one worker from the current workforce level.

The instance 1131111 has the greatest amount of scheduled overtime, with the 'AW' version of the instance (the collective stage one model) being the only cause of this; a maximum of 30 overtime hours are scheduled with a workforce of size 14 in this case with the problem becoming infeasible once a further worker is removed and the workforce size decreases to 13 clinic-based workers, while the NW1131111 instance (independent stage one model) has zero scheduled overtime until it becomes infeasible at a workforce size of 12. Although both of these instances have the same number of scheduled clinics (15 in total), the collective model instance requires a minimum of 87 shifts for CCA/CSAs while the independent model instance requires a minimum of only 69 shifts for CCA/CSAs. This is due to more trailer clinics being scheduled in the AWNW1131111 instance (a total of 11) which only require three CCA/CSAs in the North West region, compared to a community clinic in this region

Figure 6.16: North West Region - Total Overtime Hours Scheduled for Comparable Instances



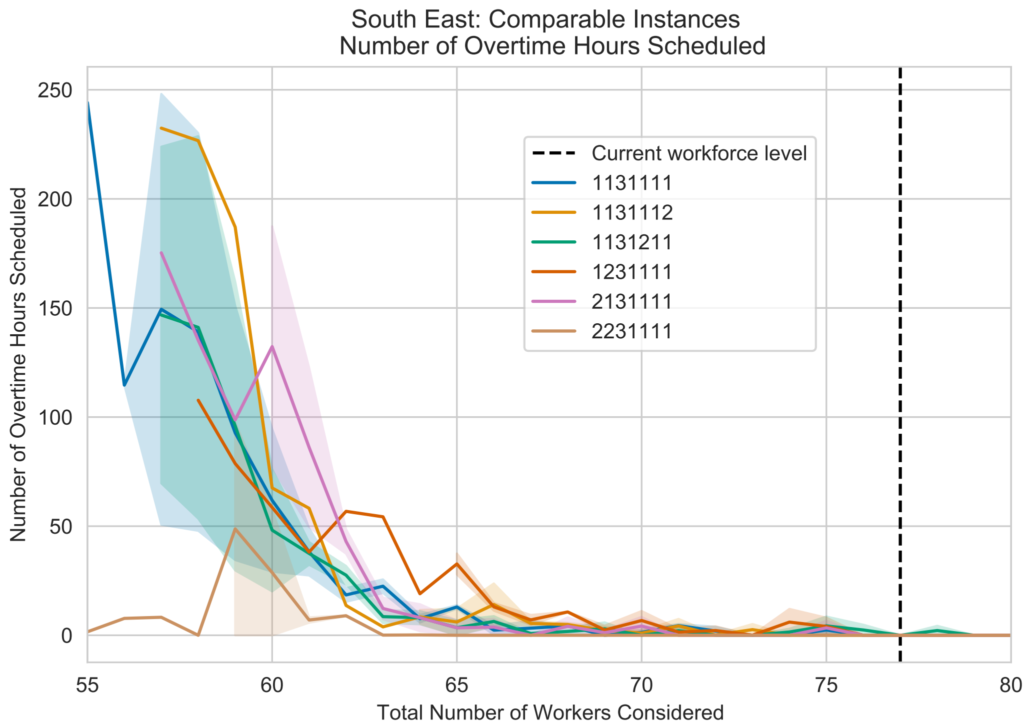
Please note: instance 1231111 remains at zero from 13 to 22 workers, while instance 2131111 remains at zero overtime hours from 14 to 22 workers

which requires a minimum of nine CCA/CSAs.

South East: For the South East region, the minimum number of workers considered where the test instance remains feasible is 55. The second-greatest number of overtime hours scheduled among the instances illustrated in Figure 6.17 is the instance where sickness is included (1131112), only exceeded by the instance 1131111 in the case of a smaller workforce. Once the workforce has decreased to 62 workers, the overtime increases significantly when sickness is included; this is due to workers having to work additional hours to cover those that have taken sick leave. The sickness rate considered in the model is typical of the WBS clinic-based workforce and demonstrates that an uplift of the workforce needs to be considered.

The instance 2131111 also has a drastic increase in overtime hours scheduled when the workforce decreases to 63 workers, then decreases when 3 more workers are removed; this is due to the problem becoming infeasible for the instance 2131111

Figure 6.17: South East Region - Total Overtime Hours Scheduled for Comparable Instances

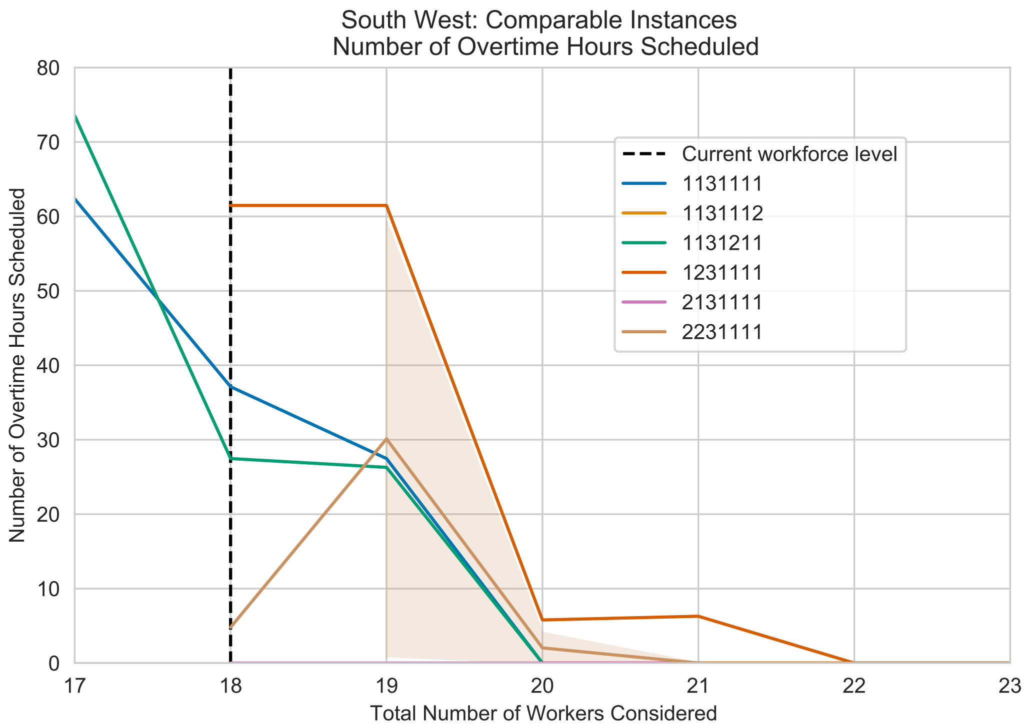


where all regions were solved collectively in stage one i.e. AWSE2131111. This is a result of larger clinics being scheduled in the South East region when it is solved collectively with others, than when it is solved independently as more of the demand is allocated to the region. Since larger clinics operate more chairs and facilitate more donors at a time, they also require more workers.

The instance with the fewest overtime hours scheduled towards the end of the iterations is 2231111; this instance is the scenario during the summer planning horizon with the constraint for the minimum number of clinics to be scheduled per region excluded. The total overtime hours scheduled begin to increase as workers are removed, similar to the trends for all other instances in Figure 6.17 though still fewer overtime hours are scheduled until the problem becomes infeasible for the AWSE2231111 case. The problem remains feasible for the SE2231111 case with overtime close to zero until the final iteration. This is due to the number of clinics scheduled in these instances, with 58 clinics in the AWSE2231111 instance and only 53 clinics scheduled in the SE2231111 instance. This demonstrates the fluctuation in required hours for the workforce over seasons, as generally fewer clinics are required to run in the summer planning horizon to meet demand and thus, fewer working hours are required.

South West: For the South West region, to ensure that all results included in Figure 6.18 are directly comparable with each other, the collective (stage one) model instances for the winter planning horizon i.e. season one are excluded due to the issue with infeasibility discussed previously. For instances 1131112 and 2131111 no overtime is scheduled, but they do become infeasible at 22 workers and 17 workers, respectively; the infeasibility of the former instance at a workforce level greater than the current workforce level at the WBS implies that this workforce level for the South West region is inadequate to handle sickness.

Figure 6.18: South West Region - Total Overtime Hours Scheduled for Comparable Instances



The instance 1231111 requires the most overtime hours while it is feasible, and this is a result of this instance having the greatest number of scheduled clinics to accommodate at a total of 16 clinics, while all other instances in Figure 6.18 have

15 or fewer.

With the exception of 2131111, all instances for the South West region at the current WBS workforce level require overtime to enable the minimum configuration of workers to be assigned to each scheduled clinic. In the case of the 2131111 instance, there are slightly fewer required shifts for Registered Nurses (RNs) which explains the lack of overtime scheduled while it remains feasible. When the workforce reaches 20 workers, a second RN is removed which is the cause of overtime for most instances at this iteration, with a third RN removed when the workforce reaches 18. This causes another significant increase in overtime hours for those that remain feasible, while the three instances are rendered infeasible at this iteration.

Comparison with WBS figures: Table 6.17 displays the number of overtime hours scheduled over the two planning horizons, both the actual WBS figures and the minimum and maximum values from the modified model with the equivalent workforce. The overtime hours are divided by north and south regions as this is how the data for the WBS was provided. Unfortunately, the data for the north regions during the winter planning horizon is not available due to the payroll being processed locally until later in 2019.

Table 6.17: Number of Overtime Hours Scheduled per Planning Horizon - Model vs. Actual WBS Figures

	Number of Overtime Hours Scheduled					
	South Regions			**North Regions**		
	Min.	**Max.**	**Actual WBS Figures**	**Min.**	**Max.**	**Actual WBS Figures**
Planning Horizon 1 07/01/2019 - 03/02/2019	0.00	63.23	517.25	0.00	28.73	N/A
Planning Horizon 2 22/07/2019 - 18/08/2019	0.00	4.80	207.25	0.00	43.05	84.75

The maximum scheduled overtime hours for both north and south are considerably lower than the actual overtime hours at the WBS for the same periods. However, for both the North West and South West regions, some instances at the current workforce level were infeasible, implying an increase in workers would be beneficial.

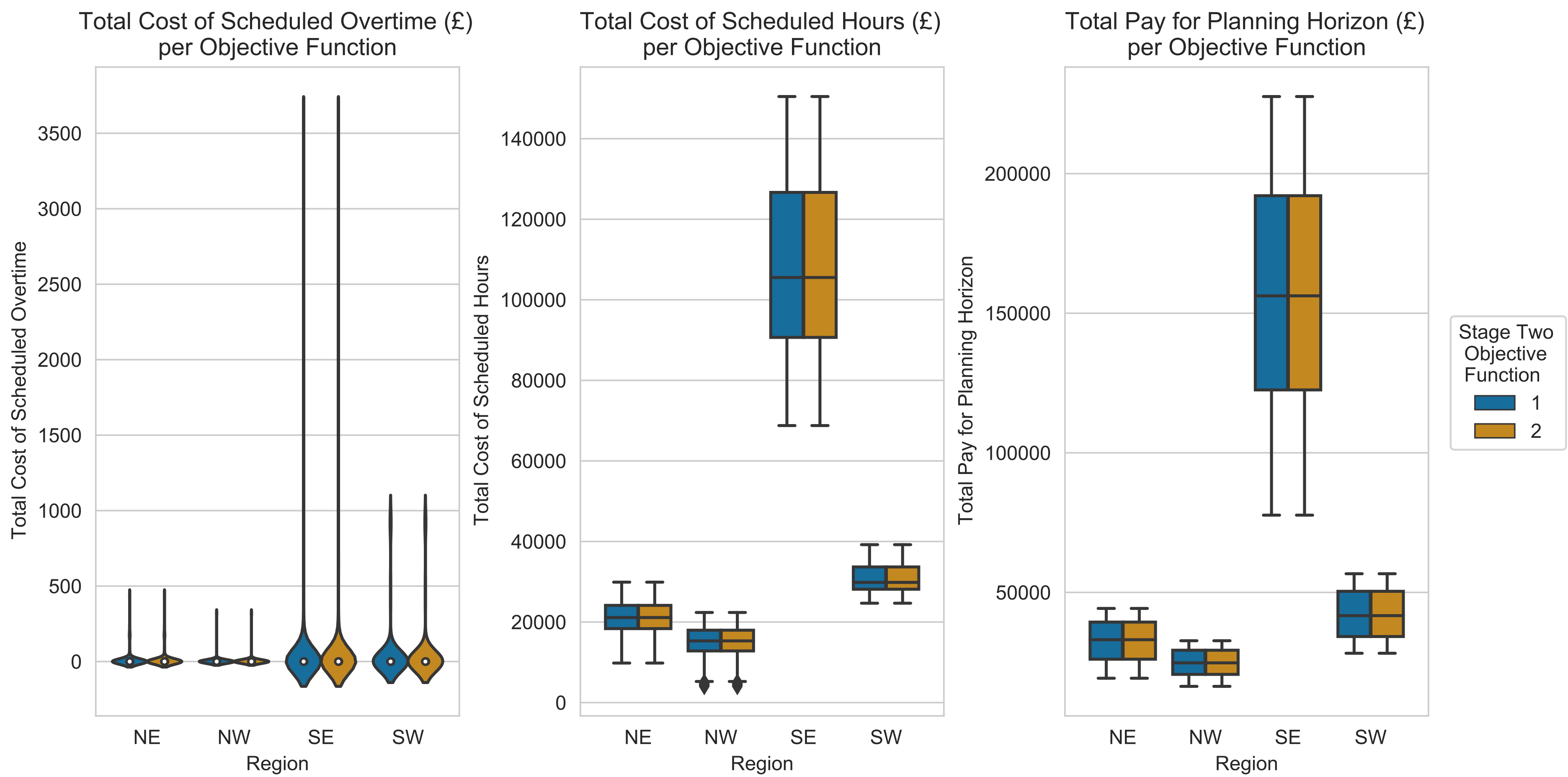
Since the 'standard' instances of 1131111 and 2131111 remain feasible for all regions (at the original modified runs), this implies that the model is capable of significantly reducing scheduled overtime for the WBS and thus, reducing monetary costs. It is worth noting that although neither planning horizon contain a bank holiday, there may be some additional 'overtime' in the form of an increased pay rate for any unsociable hours worked i.e. before 6 a.m. and after 8 p.m., though these numbers would be minimial since clinics that require working at these times are not common.

Objective Function

Two alternative objective functions are considered in the BCWSM; the minimisation of total cost of scheduled overtime (4.34) and the minimisation of the total cost of scheduled hours (4.35). Both of these objective functions also consider the minimisation of weekly overtime and undertime, and the penalty function to discourage the model from scheduling Deputy Supervisors for the clinic role of Supervisor, unless required.

Figure 6.19 displays the total costs for overtime and scheduled hours and total pay over the whole planning horizon per region, separated by objective function to allow for comparison. For all types of costs, the performance of both objective functions is identical across all regions, as shown in Figure 6.19.

Figure 6.19: Total Costs per Objective Function for Comparable Instances



Only instances that were run for both objective functions are included i.e. instances with alternative parameters in stage two are excluded.

Workforce Utilisation

The percentage of contracted hours typically available to be utilised at a blood donation clinic varies per region as a result of clinics restricted to weekdays only in the North West and South West regions, whilst clinics may occur on a Saturday in the North East and South East regions. For the North West and South West regions, the estimated amount of contracted hours that are used as annual leave or training is 13.5%, with this figure for the remaining regions estimated to be 11.6% of contracted hours. With the inclusion of the average sickness rate of 6%, this takes the figures to a total of 19.5% for the North West and South West regions and 17.6% for the North East and South East regions.

We define a workforce utilisation rate to be the percentage of contracted hours that are utilised as clinic hours. Due to the variable nature of a mobile blood donation clinic model, along with the estimated hours that will be used as annual leave, training or sickness leave, the WBS is unlikely to achieve a clinic-based workforce utilisation rate greater than 80% without scheduled overtime.

Figure 6.20 presents the utilisation rate of the whole workforce per region over comparable instances, to observe the direct impact of variables between instances on the utilisation of workers. For both the North East and North West regions, the instances without an enforced minimum number of clinics scheduled in stage one (i.e. instances 1231111 and 2231111) have significantly lower utilisation rates due to the reduction of scheduled clinics over the planning horizon. This implies that if the WBS were to adopt the collection strategy of reducing the minimum number of clinics scheduled in the north regions each planning horizon, the workforce in each region would require a reduction in size and/or contracted hours.

For the standard instances (1131111 and 2131111) where the clinic schedule generated in stage one is representative of the WBS collection strategy pre-COVID-19, the utilisation rates at the current workforce level is approximately 75% for the North East region, which is close to the ‘target’ of 80%. However, as Figure 6.15

shows, at the current workforce level for these instances, up to 15 hours of overtime are scheduled. For the North West and South East regions, the utilisation rate at these standard instances in Figure 6.20 are 67% and 64-67% for each region, respectively. The 2131111 instance for the North West region (in the comparable case) becomes infeasible at the point that the current workforce is reached. These utilisation rates align with zero overtime (or infeasibility in some North West instances) as displayed in Figures 6.17 and 6.16 which infers that the configuration of the workforce (in terms of skill mix and contracted hours) in each region would benefit from change, to promote the utilisation of contracted hours and continue to minimise overtime.

Conversely, for the standard instances in the South West region (in the comparable case), the utilisation rate of the workforce is 83% for the 1131111 instance and an average of 77% (over region independent and collective instances) for the 2131111 instance. These utilisation rates align with 29–38 hours of overtime, as displayed in Figure 6.18 which implies that the workforce should be increased in size and/or contracted hours to prevent overtime whilst maintaining a high utilisation rate. Please note that the line plot for the 1131111 instance for the South West region in Figure 6.20 is identical to that of instance 1131211. Additionally, there is no included data in Figure 6.20 for the collective (stage one) model instances for the South West region during the winter planning horizon (season one) due to the issues with infeasibility.

Figure 6.20: Workforce Utilisation Rates per Region for Comparable Instances

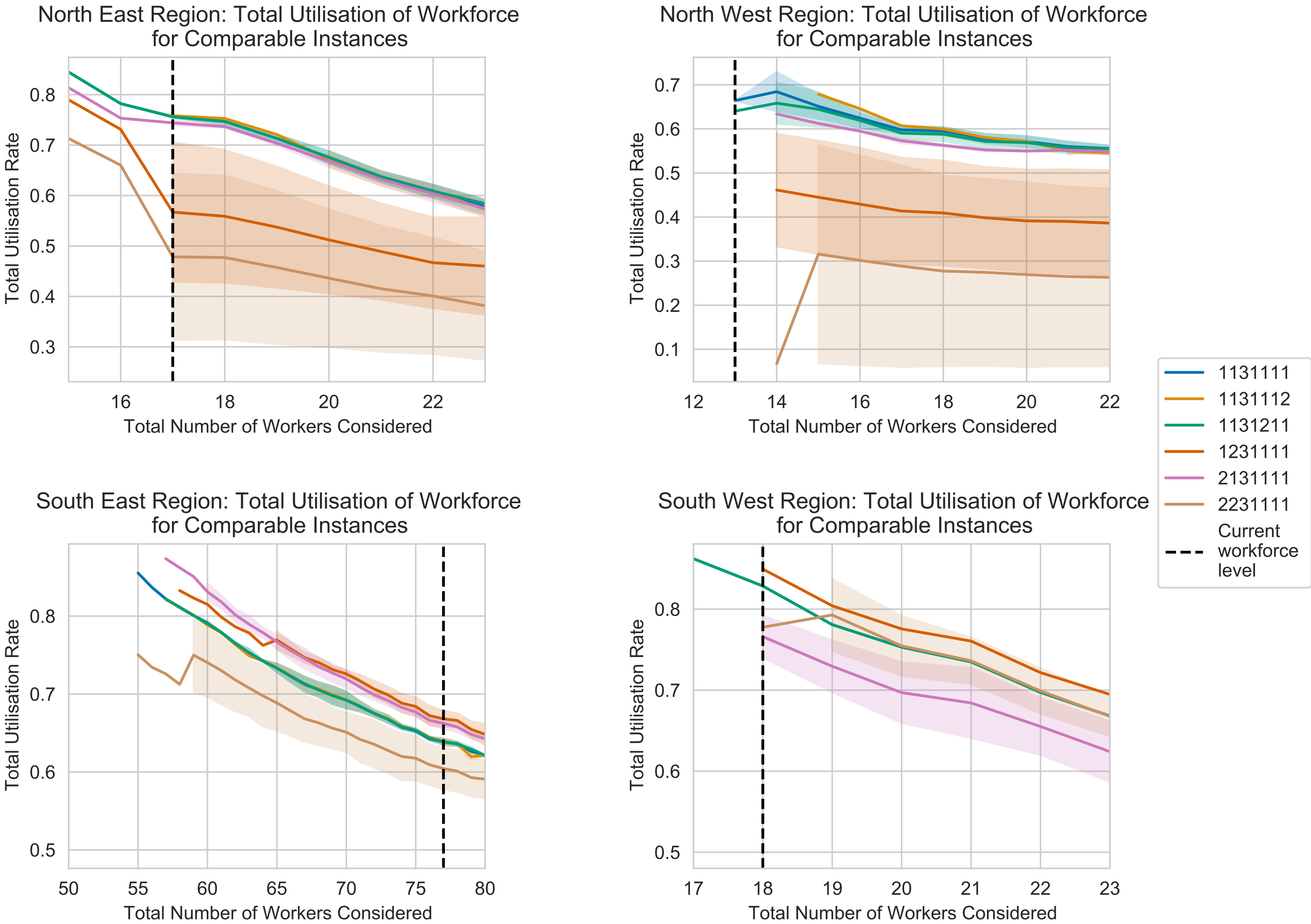


Figure 6.21: Workforce Utilisation Rates per Role per Region for 1131111 Instance

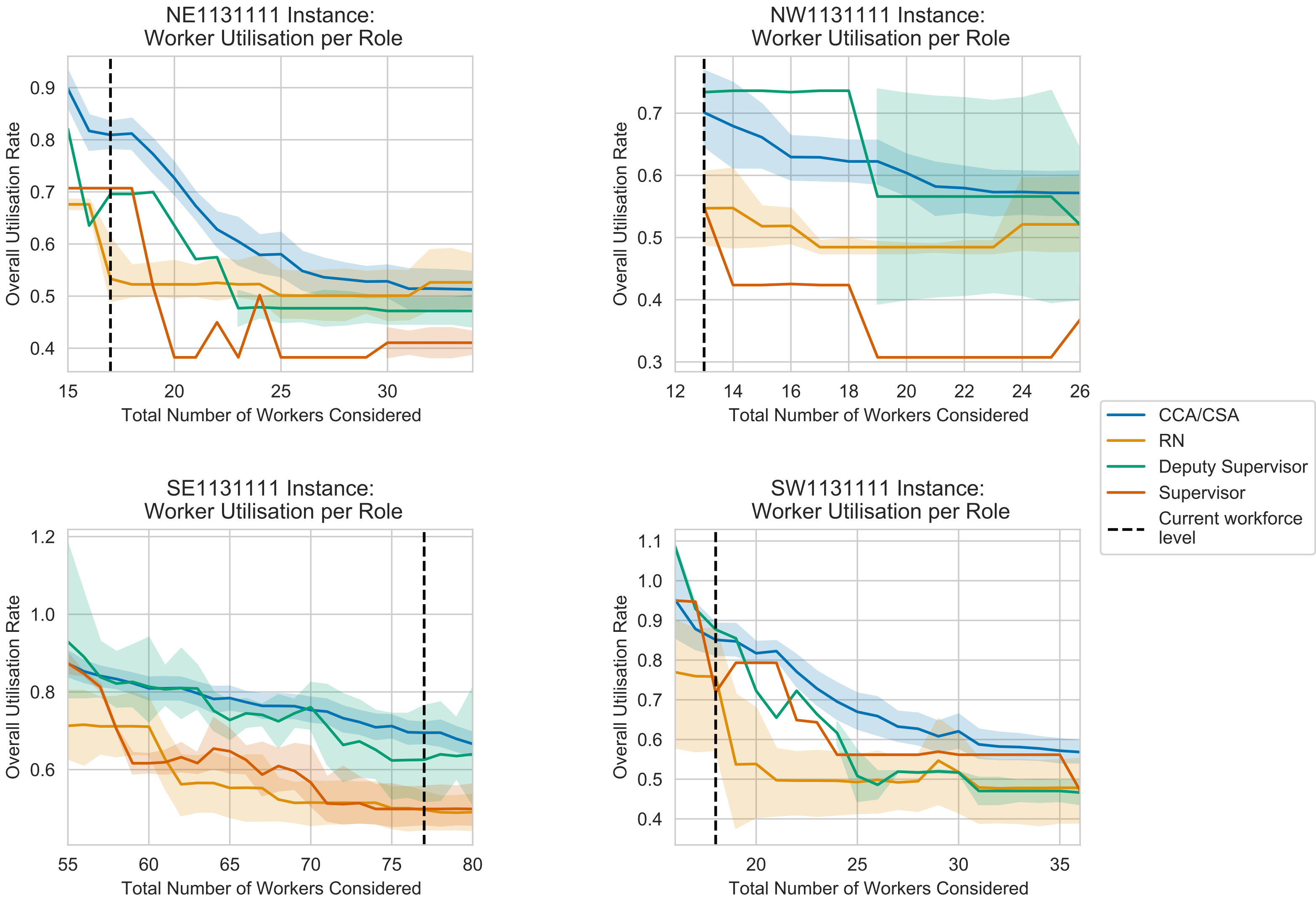


Figure 6.21 presents the utilisation rates per worker role per region in the case of the 1131111 instances, solved independently of other regions during stage one. Since dummy workers are removed before actual WBS workers in this instance until this prevents the problem from becoming infeasible, at the current workforce level, the workforce is likely to be equivalent to the actual workforce at the WBS at the time that the data was shared. For all regions in this instance, the utilisation of Registered Nurses (RNs) is the lowest of all roles at the current workforce level - although this ranking is shared with Supervisors for the North West and South West regions. For the North West and South East regions, the mean utilisation rate of RNs at the current workforce level is approximately 55%; this indicates that the number of RNs in these regions may be greater than required. The South East region has 16 RNs and thus could indeed reduce this number to reduce costs and drive efficiency, however the North West region has only two RNs and requires this as a minimum to allow clinics to continue to be operated when one RN is on leave.

Figure 6.21 implies that the number of CCA/CSAs in the South West region needs to be increased to manage resource demand without over-utilisation (i.e. relying on overtime to alleviate short-staffing situations as a result of annual leave or sickness). Since Deputy Supervisors are assigned to the clinic role of CCA when a Supervisor is available, the high utilisation rate for Deputy Supervisors in the South West region supports this conclusion. With a mean utilisation rate of 81% for the current WBS level of CCAs and CSAs in the North East region in Figure 6.21, it appears that the region would benefit from an additional CCA/CSA to lessen the pressure on the current workers and to reduce overtime.

6.3.3 Summary: Blood Collection Workforce Scheduling Model Results

This section has presented both the computational results and solutions to the Blood Collection Workforce Scheduling Problem from the model developed in Python, with all four regions solved independently of each other.

It is evident that the alternative objective functions perform identically in regards to minimisation of workforce-related costs such as cost of scheduled overtime and total pay over a given planning horizon. Both objective functions also perform similarly computationally, with no noticeable difference between run times or time taken to reach an optimal or near optimal solution due to the optimality gap. Therefore, either objective function could be utilised by the WBS; however, since the main goal of the WBS is to reduce overtime, it makes sense for objective function one to be the objective function of choice.

The model can generate an optimal (or near optimal) workforce schedule for the three 'smaller' regions – in terms of problem size – namely the North East, North West and South West regions in less than 10 minutes per region. For the much larger problem for the South East region, the total run time increases to a maximum of approximately 12 hours in cases where the workforce is reduced by at least 15 workers from the current workforce level. However, at the current workforce level (and up to a reduction of 10 workers for most instances) the model generates a (near) optimal schedule in less than three hours. With workforce schedules for all four regions collectively produced in under four hours by the BCWSM, this is a drastic increase in efficiency compared to the current practice, where a total of approximately 48 hours are required to manually create a four-week schedule. It is worth noting that the time taken to input the annual leave, training and sickness leave (if the latter is known at the time of scheduling) would currently need to be manually input before the model can run, which would need to be considered in addition to these figures. However, with the schedules being automated, this still significantly reduces the manual work for the workforce planning team, where additional adjustments to the schedule may also be made, as required. This is one way in which we have answered research question two: 'How can mathematical modelling help to schedule the clinic-based workforce at the Welsh Blood Service?'

As a result of assessing the workforce utilisation rates and scheduled overtime hours at various staffing levels for all regions, it appears that the South West region would benefit from an increase in workers, particularly of the role of CCA/CSA. Since all of the winter planning horizon instances with the collective stage one model input are infeasible from the first instance, and overtime is required at the current workforce level for all feasible instances, we can assume that the workforce for the South West region may require a review by the WBS if the collection strategy returns to the pre-pandemic model. Conversely, the low utilisation rate of RNs in the South East region and near-zero overtime hours at the current workforce level (and for up to ten fewer workers thereafter) implies that costs could be reduced for the WBS by a decrease in the number of RNs in the region's workforce. The ideal workforce sizes and mixes for both the of the north regions depend on the collection strategy i.e. how clinics are scheduled in stage one; with the results from the 'standard' instances that are representative of the WBS pre-COVID-19 strategy implying that both the North East and North West regions could benefit from an additional worker (likely a CCA/CSA) to cope with annual leave and sickness, and reduce overtime hours from being regularly required. However, if the WBS choose to operate slightly fewer clinics in the less efficient north regions as selected by the collective stage one model, the current workforce levels appear adequate. These insights have the potential to support the WBS in their decision making regarding workforce planning and to save monetary costs where possible.

Finally, the BCWSM provides the potential for the WBS to reduce the overtime hours required over a given planning horizon, as evident in Table 6.17, especially when the model is used in conjunction with stage one i.e. the BDCSM. If the workforce level insights discussed above are utilised to improve each region's team, then this would drive a further reduction of the overtime hours required to collect enough blood donations to meet the demand, and allow regions to cope more effectively with sickness of clinic-based workers. However, the balance between additional contracted hours (for new workers or an increase in contracted hours for existing workers) and the cost of overtime is complex and requires further inspection by the Welsh Blood Service to keep costs to a minimum and to ensure overtime is not too heavily depended on to meet the blood product demand.

The following chapter, Chapter 7 concludes this thesis, with a summary of each prior chapter, the findings from our research and how we have answered the three research questions presented in Section 1.5. Additionally, we discuss the limitations of our models and how this project has been impacted by the COVID-19 pandemic, followed by areas of interest for further work related to this research.

Chapter 7

Discussion and Conclusion

This final chapter summarises the content of this thesis, identifying contributions and how the research questions listed in Section 1.5 have been answered. Additionally, to set the thesis into context, a discussion surrounding assumptions, implementation, scalability and robustness of the two-stage model is presented in Section 7.3. Subsequently, a concluding statement is provided to close the thesis.

7.1 Research Summary

Chapter 1 introduced the history and importance of human blood as a life-saving resource, and the typical layout of a blood service's supply chain and its multiple echelons. The Welsh Blood Service (WBS) was also introduced and their requirements for this research as an industry partner under the KESS 2 scholarship scheme. The three following research questions were introduced, focussing on increasing efficiency of the collections process for the Welsh Blood Service:

1. How can mathematical modelling help to schedule the Welsh Blood Service's blood donation clinics more efficiently?

2. How can mathematical modelling help to schedule the clinic-based workforce

at the Welsh Blood Service?

3. Can these mathematical methods be integrated into a decision support tool for planners at the Welsh Blood Service to use?

Chapter 2 provided a detailed review of existing literature relevant to this thesis i.e. publications that consider optimisation and/or modelling of the collections echelon in the blood supply chain. Gaps in the literature were discussed in the summary section of this chapter, such as the lack of research into both optimal workforce scheduling for blood donation clinics, and matching blood supply to demand. This identifies a need to focus on these topics, which are addressed in Chapters 3 and 4.

Chapter 3 addresses research question one; the current (pre-COVID-19) collection model of the WBS is presented, alongside restrictions and limitations of this practice, such as the tedious manual process of creating a clinic schedule and potential overcollection of blood donations. Following this, the Blood Donation Clinic Scheduling Problem (BDCSP) was formulated as a linear programme to optimally schedule clinics over a four-week planning horizon. Three alternative objective functions were presented: the minimisation of the number of clinics scheduled, the minimisation of the number of units of donated blood that are collected in excess of the demand (i.e. overcollection), and the minimisation of both. Multiple constraints are presented to ensure that the output schedule of the model is as realistic as possible and is a feasible solution for the WBS.

Chapter 4 addresses research question two; the current practice at the WBS for scheduling clinic-based workers and associated restrictions are presented and the limitations of this process were described. These limitations include inflexibility of the clinic teams, regular use of considerable overtime, and the time-consuming manual process of creating a schedule. Subsequently, the Blood Collection Workforce Scheduling Problem (BCWSP) is presented, also formulated as a linear programme, to optimally assign clinic-based workers to the clinics determined in the model detailed in Chapter 3. This formulation utilises modes to select a clinic role and a driving role for each worker for each clinic, and considers all workers within a given region in replacement of predetermined teams. In the original model formulation, two alternative objective functions were introduced: the minimisation of overtime cost and the minimisation of the total cost of scheduled hours. A modified version of this model is presented to ensure the output is realistic for the WBS (to more fairly distribute clinic hours across workers). It includes an improved version of each objective function which consider the minimisation of the total weekly overtime hours and total weekly 'undertime' hours per worker, and the addition of a new constraint to promote the distribution of working hours to all workers.

Chapter 5 addresses research question three with the description of the development of a prototype decision support tool for scheduling blood donation clinics at the Welsh Blood Service using Excel and OpenSolver, in addition to the development of a Blood Donation Clinic Scheduling Model (BDCSM) and a Blood Collection Workforce Scheduling Model (BCWSM), both built in Python and utilising PuLP. The aims of these models are discussed, alongside how data is input and utilised. The design of experimental scenarios is described for test instances for both the BDCSM and BCWSM developed in Python, with the motivation behind these design decisions presented.

Chapter 6 presents the experimental results of the designed test instances for the BDCSM and the BCWSM, including the computational results and the solutions. The BDCSM utilising objective function three (the minimisation of both the number of clinic days scheduled and estimted overcollection over the given planning horizon) performs well with significantly less overcollection and fewer clinics (up to as many as 20 fewer) scheduled than the actual WBS figures for the same planning horizons, thus reducing clinic-associated costs and potential wastage of blood products. Additionally, the experimental investigation has revealed that an optimal four-week clinic schedule is generated by the model in significantly less time than it requires the clinic planning team at the WBS to create an initial schedule; with the potential to reduce the process from an average of five working days (current practice at the WBS) to approximately eight hours. Collectively, these results demonstrate how the BDCSM could increase efficiency of the clinic scheduling process, reducing costs, wastage and time, thus answering research question one.

The results of the BCWSM illustrate how the model (when used in conjunction with the BDCSM) is capable of reducing scheduled overtime of clinic-based workers, by as much as 400 hours over a four-week planning horizon, and increases utilisation of contracted hours to minimise costs associated with overtime. Additionally, the results provide some insight regarding the staffing levels per region and how these contribute towards overtime or unnecessary costs for the WBS. Finally, the BCWSM generates a (near) optimal workforce schedule for a four-week planning horizon in significantly less time than it takes the workforce planning team at the WBS, with the potential to reduce the process from an average of 48 hours to only four hours. This answers research question two, with the BCWSM having the potential to increase the efficiency of the scheduling of the clinic-based workers in terms of both cost and time.

7.2 Contributions

One of the main contributions of this research is the extensive literature review provided by Chapter 2. This review details all existing publications that consider optimisation and/or modelling of the blood collection echelon of the blood supply chain. A detailed taxonomy was developed to present a clear categorisation of existing research, with an identification of areas that require further research. This review focusses specifically on the blood donation collection process, where this is a distinct lack of existing reviews, and provides an up-to-date evaluation of this particular field of research, with a prior version of this literature review published in 2020 [101].

Specific gaps in the existing literature were identified in this review, namely the optimisation of workforce scheduling for blood donation clinics, and the direct matching of blood supply to demand for blood products. The Blood Donation Clinic Scheduling Model and the Blood Collection Workforce Scheduling Model presented in Chapters 3 and 4 respectively, with results discussed in Chapter 6, fill these gaps by contributing mathematical models to optimise the scheduling of both mobile blood donation clinics (by matching supply to demand) and the clinic-based workforce.

7.3 Discussion

To set the research presented in this thesis into context, this section addresses various aspects of the two-stage model and its potential role at the Welsh Blood Service and elsewhere.

7.3.1 Assumptions and Limitations

In order to create a two-stage mathematical model, some assumptions and simplifications were required to be made. These will be discussed per individual stage of the model.

The Blood Donation Clinic Scheduling Model

Regarding the first stage of the model, the BDCSM, the following assumptions were made:

- It is assumed that the demand is known. As discussed in Section 3.1.8, the WBS is undertaking research to identify the 'true demand', as supplied hospitals work independently from the WBS and are in control of determining their own blood product orders. This means that at present, the actual demand is not known as the hospitals do not report back their usage. However, throughout this research, we have assumed the demand to be that of meeting the hospitals' orders. The Business Intelligence department at the WBS has been working on improving demand forecasting, and thus we also assume that if the model were to be implemented at the WBS, then these demand forecasts would be input into the model with some tolerance factored into the figures. This tolerance is to ensure that if the number of blood donations collected during a given week differs from the estimated supply at the time of the schedule being generated, then there would still likely be sufficient blood product available to cope with demand.

- It is assumed that the estimated supply data that would be input into the model, if implemented at the WBS, is a good prediction of viable collections expected from each scheduled clinic, i.e. that the actual viable collections from a given clinic is not likely to vary greatly from the forecasted viable collections. Similarly to the demand forecasts, the Business Intelligence department at the WBS has also been improving their estimated supply forecasts. There is still uncertainty involved in donor attendance and blood donation supply which will be discussed further in Section 7.3.5. For the experimental results, we assumed that the estimated supply per day per clinic is the same for any day that the clinic operates during the planning horizon. However, the supply forecasts that the WBS have been working on do not assume this and likely incorporate variation between days. The possible implementation of this is discussed in Section 7.3.3.

- It is assumed that the WBS are to decide the weights assigned to the terms in the objective function if objective function three (3.5) is utilised in an implemented model, i.e. that weights will be assigned to both the number of clinic days scheduled and the estimated overcollection over a given planning horizon. This would allow the WBS to establish how they value each component, and how they should be compared to each other.

- The current BDCSM that is presented in this thesis assumes that all clinics are to be scheduled at the same time. In practice at the WBS, the clinics that are considered to be 'hard to book' such as clinic tours, are often booked further in advance, and then the remaining clinics are scheduled at a later date. The rationale behind the decision to schedule all clinics over a given planning horizon at the same time is that since the generation of an initial clinic schedule would be automated (saving a significant number of hours of manual scheduling), all clinics can be scheduled as far in advance as desired. Scheduling 'hard to book' clinics further in advance than other clinics however could be easily incorporated into the model, and this is discussed in Section 7.3.3.

- The current model developed in Python is based on the planning horizon being four weeks since this is how the WBS schedule their clinics. However, if this were to change to a greater number of weeks such as eight or 16 weeks, the model may no longer run in a reasonable time, and may need to utilise heuristic methods for the case where all regions are solved collectively. The use of heuristic methods for scalability is discussed further in Section 7.3.4.

In addition to the above assumptions, there are some limitations associated with the BDCSM. The main script contains the various constraints and objective functions that were presented in Section 3.3. If the collection strategy were to change, such that the constraints are required to be changed, then the WBS would need to implement this. For example, in cases where perhaps the maximum number of clinics permitted to be scheduled per day per region may change, this is a relatively straightforward change to the main script, only changing the number on the one side of a constraint. However, if a new constraint needs to be included, this would require members of the IT department at the WBS to have in-depth knowledge and understanding of the mathematical constraints and how to write new constraints, as required. Similarly, if a new clinic duration pattern or availability were to arise, then this would need to be written into the model by the WBS. This is discussed further in Section 7.3.3.

The Blood Collection Workforce Scheduling Model

Regarding the second stage of the model, the BCWSM, the following assumptions were made:

- It is assumed that a given worker is only assigned to work at clinics that are based in the same associated region as the worker. In practice at the WBS, workers may be assigned to work a clinic in a different region (if the worker is willing to) in cases where there is a staff shortage. However, since this is not a particularly common occurrence and is often the result of short-notice operational offline decisions made after the time that a workforce schedule is published, for simplification purposes, we choose not to consider these cases in the model. This also allows for regions to be solved independently from each other and reduces the complexity of the problem.

- It is assumed that the length of working day for a given clinic is true. In practice, the length of working day for a clinic may vary slightly from the provided number in the clinic data due to factors such as traffic, the closure of a clinic delayed due to donations still taking place, breaking down of a vehicle, issues with the clinic venue, etc. This means that the estimated number of hours assigned to workers by the model is likely to vary somewhat from the actual number worked.

- It is assumed that any scheduled overtime is paid additionally to a worker's salary, and is not considered as time off in lieu. The WBS generally pay overtime in addition to salary, with an increased pay rate for hours worked in excess of 'full-time hours' i.e. over 150 hours per four-week period. However, according to the workforce data provided by the WBS, one clinic-based worker instead receives overtime as additional time off in lieu. Since this is only one case and the additional leave is unlikely to have an impact, we simplify the model and consider only paid overtime.

- For the clinic-based workforce in the South East region, there are currently four set teams at the WBS. However, for the model, we assume that this region is instead one flexible pool of workers that may all work with any other group of workers in the region. We make this assumption since it is a goal of the WBS to move towards a more flexible staffing model in this region to promote efficiency and to avoid situations where an excess of workers are assigned to a given clinic.

- It is assumed that each worker has 30 days of annual leave. Since the WBS were not able to provide each worker's annual leave allowance or duration of employment for anonymity reasons, we assume that all workers have the mid-point annual leave allowance of 30 days, as annual leave for NHS employees ranges from 27 to 33 days dependent on the duration of employment. This assumption is made for simplification purposes to allow leave schedules to be generated for the experimental results.

- Similarly, since the WBS could only provide the salary band of each worker and not the actual salary (or their duration of employment), we assume that a worker's salary is the mid-point of their salary band. The salaries per band for NHS organisations is widely accessible online and is discussed in Section 4.1.3.

- Similarly to the BDCSM, the current model developed in Python is based on the planning horizon being four weeks since this is how the WBS schedule their workforce during 'business as usual'. If this were to change to a greater number of weeks, the model may no longer run in a reasonable time, and may need to utilise heuristic methods for the case where all regions are solved collectively. However, due to the nature of workforce scheduling and considerations needing to be made regarding annual leave etc. it is unlikely that the WBS would increase the planning horizon for workforce scheduling. During the COVID-19 pandemic, workforce schedules have been created for one-week periods at a time to better cope with increased absences due to sick leave and isolation periods of workers. If the model were to only consider a one-week planning horizon then it would not be able to minimise scheduled overtime effectively, since this is calculated over four-week periods, and thus is best utilised over four-week planning horizons to maximise the benefits of its use for the WBS.

Alongside these assumptions, there are limitations associated with the BCWSM. Similarly to the limitations described for the BDCSM, any changes in the staffing model for the WBS regarding constraints or availability patterns etc. would need to be manually changed in the main script by the IT department at the WBS.

7.3.2 Purpose and Use of Decision Support Tool

A decision support tool or system is defined as a 'computer-based interactive system that supports decision-makers rather than replaces them; utilizes data and models; solves problems with varying degrees of structure' [32]. The initial aim of this project was to create a mathematical model which could be embedded into a decision support tool for the WBS to improve how they schedule their clinics and clinic-based workforce. Utilising the initial schedules produced by the two-stage model as a starting point for the planning team would allow for potential reductions in both monetary costs and overcollection of blood donations for the WBS.

If the two-stage model were to be implemented at the WBS, in the form of a decision support tool, it would provide a close to optimal initial schedule (for both clinics and the clinic-based workforce) that is generated in significantly less time and with a reduction in manual work required. The schedule, either clinic or workforce, can then be easily adjusted as required to account for factors such as a scheduled clinic being newly unavailable, or a worker requesting annual leave or taking sick leave etc. It would therefore be helping to inform planning decisions, not explicitly making these planning decisions, to motivate the creation of more tactical schedules that align with the goals of the WBS.

Both the BDCSM and the BCWSM have been developed to create various outputs, usually in CSV format since this is compatible with various computer operating systems. There are multiple versions of schedules, which are described in Section 7.3.3. These are currently output as CSV files but could easily be adjusted to be in Excel format, as this is familiar to the users of the tool, or as part of the development of a GUI, if implemented.

7.3.3 Implementation and the Impact of COVID-19

For the two-stage model to be implemented at the WBS, the staff in the IT department at the WBS would need to be heavily involved. During discussions that took place early in the project with senior stakeholders at the WBS and members of their IT department, it was agreed that some capacity would be provided by the IT department to develop a graphical user interface for the two stages of the model, with the possibility for further capacity to be allocated to implement the models as a decision support tool, provided the value of the tool was demonstrated once the models were more fully developed. Unfortunately, due to the impact of COVID-19, the collection model of the WBS has changed considerably since March 2020 and at present, there is no plan to return to the previous collection model given the ongoing limitations associated with the pandemic. The collection strategy has moved from aiming to provide a local service to all donors across Wales, to utilising clinic venues that are large to enable social distancing. These large venues that previously were only one-day clinics are now scheduled for multiple consecutive days and a larger donor panel is invited to the clinic, effectively replacing the more local clinics for many donors. Trailer clinics are no longer being operated as they are too much of a confined space to allow for social distancing between donors.

This new collection strategy only affects stage one of the model (BDCSM) since the assignment of workers to clinics remains the same, and the new strategy could be relatively easily implemented into the model by altering the input data such as clinic duration patterns and frequency-based availability, and altering the clinic venue related constraints to prevent any trailer clinics from being scheduled. The new strategy may leave little opportunity for improvement through use of the BDCSM as decision support during the clinic scheduling stage as the problem is now much more constrained. Since stage two of the model (BCWSM) utilises the output of stage one, all collection strategy changes are effectively implemented through the input of a clinic schedule, and this model could be used to improve efficiency of workforce scheduling for clinic-based workers at the WBS.

The COVID-19 pandemic has placed increasing demand on the IT and Business Intelligence departments at the WBS due to necessary blood product demand and donor attendance forecasting efforts, in addition to devising a new collection strategy and locating the best clinics for social distancing. As a result of this, there is currently no capacity to develop a graphical user interface for the two Python models (the BDCSM and the BCWSM) and therefore, in the near future at least, no capacity for implementation of the models.

Client Feedback

In February 2022, Jayne Davey, Interim Blood Supply Chain Lead of Collection Services at the Welsh Blood Service provided the following client feedback:

> *The ability of a blood service to ensure a robust supply chain relies heavily on the strength of its relationship with donors. At the beginning of the project, due to the long-established ways of working at the Welsh Blood Service, there were a significant amount of constraints to be included in the clinic scheduling process. These included working practices around staffing, extensive community presence at low populated communities and a strong tradition of taking the service as close to the donor as possible. This resulted in limitations on how far the efficiency of the collections process could be improved.*
>
> *The requirement for social distancing associated with the SARS-*

CoV-2 coronavirus pandemic (COVID-19) turned the whole blood collection portfolio on its head and signalled the loss of the Welsh Blood Service's mobile donation clinic trailers, all of the university and company clinics and most of the small community venues. It necessitated a radical change in our collection strategy and as a result, donors have adapted their behaviours and expectations of the service. This temporary collection strategy is unsustainable in a post pandemic world and so, further change will be needed that promotes the development of a sustainable, future proof blood collection clinic portfolio that is able to meet patient demand for blood and blood components, values the donor as the vital supplier of this life saving commodity and aligns to expectations of Wales as it strives to create a better world for our future generations. This stabilisation requirement provides the Welsh Blood Service with the opportunity to improve our collection clinic model which in turn enables us to incorporate greater flexibility and optimisation of clinic and workforce scheduling in the future, increasing the potential impact of the two-stage model developed during this project.

Requirements for Implementation

From Jayne Davey's feedback it is evident that there is still support for the model to be implemented in the future. In the following, we discuss the requirements for this.

Early in the project, the senior stakeholders and members of the IT department at the WBS agreed to allocate some capacity to make the mathematical models useable, in the form of a decision support tool. For this to happen, a graphical user interface (GUI) would need to be developed to allow the end-user i.e. clinic or workforce planner to interact with the models. There is currently a digital rostering tool (to manually create schedules for clinic-based workers) that is implemented at the WBS that was created by their own IT department. It is possible that the two-stage model could be integrated into this GUI, but particularly the second stage of the model, the BCWSM. Otherwise, the first stage (the BDCSM) could be a separate tool, as the employees who schedule the donation clinics are independent from the workforce planners, using a similar GUI to the existing digital rostering tool.

The current output of the models are schedules in a CSV format. For the BDCSM, there is a binary schedule output that is saved each time it runs, with the dates of the planning horizon as column headers, the unique clinic keys as rows, and binary values in each 'cell' to denote whether a clinic is scheduled on a given day, or not. There is also the alternative schedule output with the estimated supply in each 'cell' in place of the binary decision variable. For the BCWSM, there are many versions of schedules saved at each run. For example, a schedule with the dates of the planning horizon as column headers, unique ID for each worker as rows, and the unique clinic key in the 'cells' for any clinics assigned to the corresponding worker on the corresponding day. Alternative schedules (with the dates still as column headers) include the unique clinic keys as rows and the unique IDs of any assigned workers, on the corresponding day, as the 'cell' values. This makes it relatively straightforward for the IT department at the WBS to take these output CSV files and display them in an accessible way for the end-user via the GUI, and/or to save the output as an Excel file for the planners to view - a format that they are already comfortable with.

In addition to the development of a GUI, both stages of the model would need to be able to communicate with the data warehouse and the WBS Blood Establishment Computer System, ePROGESA, which are both hosted on the servers in the WBS data centre. It has been indicated by the IT department that it would be straightforward to store a decision support tool on the same servers, that could be run remotely from PCs by the clinic and workforce planners, and that the tool could retrieve the required data from elsewhere on the servers. Ongoing maintenance of the tool(s) would be required, to fix any bugs and to make any changes to the main script of the models, as suggested in the discussion surrounding the limitations of the models in Section 7.3.1. For the tool(s) to be used long-term, any changes to the collection strategy or staffing strategy at the WBS would need to be included in the Python scripts of the models. This requires members of the tool maintenance team to have a working knowledge of the mathematics involved, in order for them to formulate any new constraints correctly. Several members of the Business Intelligence team at the WBS are familiar with OR methods, including linear programming, and therefore it would be beneficial if this department could be involved. If implementation were to happen in future, several knowledge transfer sessions with the IT and Business Intelligence departments, would need to take place.

At present, due to changes to the collection strategy as a result of COVID-19, the current first stage of the model, the BDCSM, is not directly applicable. The BD-CSM, would require minor adjustments before implementation could begin, which mainly consist of additional clinic duration patterns and changing the input clinic data to ensure no temporarily unavailable clinics are scheduled. Additionally, to more closely reflect the clinic planning process at the WBS, any already scheduled 'hard to book' clinics (which are typically booked much further in advance of other clinics) could be read into the model via and Excel or CSV file with the dates that they are scheduled. With the addition of a few lines of code for the input of booked clinics and a new constraint, it would be relatively easy for the BDCSM to consider these clinics and schedule other clinics around them for the remainder of the planning horizon. The consideration of estimated supply per specific day, as discussed in Section 7.3.1, would also be relatively simple to include in the model by adding a 'day' component to the estimated supply variable in the main script, and ensuring that the input clinic data includes the relevant data.

If the two-stage model were to be implemented in future, the process would likely consist of a development stage (to create a GUI and to integrate the tool with existing systems) followed by a trial stage with the end-users to gain feedback on usability. For the successful implementation of a decision support tool, it is imperative to closely involve the end-users in the process to maximise the usability and accessibility of the tool. Usability testing with end-users may take various forms, but a common method, and perhaps the most suitable for this particular scenario, is remote testing. This method involves video recording computer screens of testers (end-users at the WBS) and audio-recording the testers as they are encouraged to 'think aloud' while interacting with the tool [55]. This collects both quantitative and qualitative data from the end-users, which could be used in combination with interviews and/or questionnaires conducted with the end-users to establish what changes should be made to the tool to improve usability.

This would likely become an iterative process, where feedback is taken on-board and the tool is further developed or adjustments made, followed by another trial by the end-users, and so on, until the tool is satisfactory, as is common practice with implementation of decision support tools [55]. This is similar to the process described for the planned pilot study of the prototype tool in Section 5.5, and may take several months to complete.

7.3.4 Generalisability and Scalability

This research project is the result of a collaboration with both Knowledge Economy Skills Scholarships (KESS 2) and the WBS, which limits the scope of the research to be focussed solely on the activities and objectives of the WBS. Consequently, the two-stage model presented in this thesis is tailored to the specifics of the (pre-COVID-19) WBS blood collection model. Since each blood service organisation across the globe has its own individual infrastructure, requirements and considerations, often unique to the local demographics and/or geography, it is unlikely that the two-stage model is directly applicable to another blood service in its current state.

The two-stage model may be applicable to blood services that rely heavily on 'mobile' clinics for their blood supply i.e. not permanent clinics that operate frequently at a fixed site, but either clinics that take place occasionally at a hired venue or in a parked trailer. An example of this would be the Scottish National Blood Transfusion Service, who operate permanent clinics at six fixed donor centres, but also rely on many mobile clinics at various locations to meet blood product demand. For the two-stage model to be implemented at another blood service with a similar collection model, there would almost certainly still be required changes to the clinic availability and duration patterns, as well as the constraints (to reflect the individual needs of the organisation) in the BDCSM. Changes to the BCWSM to account for different working day patterns would also likely be required, as a minimum.

In addition to changes required in the problem formulation, the exact method of integer linear programming is likely not suitable for much larger and more complex scheduling problems. Stage one of the model, the BDCSM, is a generalisation of the proven NP-hard job-shop scheduling problem, whilst stage two of the model, the BCWSM, is an extension of the multi-mode resource-constrained project scheduling problem, which is also a generalisation of the NP-hard job-shop scheduling problem [53]. Since Wales is a relatively small nation with a small blood collection model by the WBS (in terms of number of clinics locations and clinic-based workers compared to other nations such as England), the test instances considered in this research have been able to be solved to near optimality within a reasonable time. However, nations with a much larger population and therefore larger demand need to operate many more clinics, and it is unlikely that this exact method can be utilised in these cases. For example, NHS Blood and Transplant, the blood service for England, operates 25 fixed clinics but also operates mobile clinics in 'thousands' of locations, according to their website [20], which also implies a significantly greater clinic-based workforce. For blood services that operate this many mobile clinics, a heuristic method is likely to be required to achieve a solution in a reasonable time due to the associated increase in complexity of the problem for both stages of the model. This is also discussed in Section 7.4.3.

Other blood service organisations, as well as the WBS in future, may also wish to incorporate statistical modelling methods to account for the uncertainty associated with blood donation clinics, to ensure that any schedule solution is robust and more prepared to cope with fluctuations in supply and demand. This is discussed further in Section 7.3.5.

7.3.5 Robustness and Resilience

There is a significant element of uncertainty involved in the blood collection process, such as donor attendance, demand fluctuation, and the more specific demand per blood type and how to collect from donors to meet this. At present, since the tactical scheduling of donation clinics happens many months in advance at the WBS, it is difficult to consider the inventory levels across blood types at the time that a schedule is produced. The WBS counteracts any low stock levels of a certain blood type by contacting donors of the blood type that live nearby upcoming scheduled clinics, to encourage them to attend. For the most part, this is an effective solution for the WBS, but in cases where it is not, they schedule 'emergency' clinics at their fixed clinic site at the WBS headquarters and again, contact local donors of the required blood type to attend. For this reason, alongside the lack of available donor demographic data due to NHS data privacy regulation, we have not considered specific blood types in the model. However, in future, it would be beneficial for the WBS to consider donor demographics during the clinic scheduling process, to enable specific targeting of a given blood type, and even to inform strategic decisions about where to locate a new clinic. This would enable the WBS to avoid the current practice of 'reactive controls' as opposed to 'proactive controls', as discussed by Mugdh et al. [64] in their framework for balancing resilience and innovation in relation to optimisation of emergency departments.

Other elements of uncertainty, such as donor attendance and demand fluctuation, are addressed through the forecasted supply and demand data that would be provided by the WBS Business Intelligence department and input into the model, if implemented. Alternatively, statistical modelling methods could be added to the model in future to consider these uncertainties and to provide a more robust clinic schedule as another potential proactive control [64]. In Section 2.2.2, it is evident that a significant proportion of papers consider both integer programming and stochastic modelling in combination to provide a robust approach [9, 25, 35, 41–43, 46, 47, 58, 79, 86, 89, 104, 105]. The majority of these papers however focus on creating a robust model for blood services that are likely to encounter disasters such as earthquakes, to ensure that there is adequate supply to deal with such emergencies. It may be beneficial for the WBS to consider these stochastic elements to develop a more resilient collection model to cope with large-scale emergencies, furthering their risk management strategy and developing a more proactive approach.

7.4 Further Work

Throughout the research detailed in this thesis, new research questions and directions of interest have arisen which are not within the scope of this research, but would further develop this research and/or would further the contributions in the field of optimisation of blood collection. These research areas are described below.

7.4.1 Pareto Analysis for Weights in Objective Functions

Objective function three (3.5) in the Blood Donation Clinic Scheduling Model (BDCSM) seeks to minimise the total number of clinic days scheduled and the total estimated supply in excess of demand for a given planning horizon. As discussed in Section 7.3.1, it is assumed that the WBS would determine appropriate weightings to assign to both of these individual terms in the objective function to prioritise accordingly between number of clinics and overcollection, since they cannot be equated. One option to determine weightings is to consider Pareto optimisation, which is commonly used in cases of multi-objective optimisation problems to ensure that any conflicting objectives are optimised simultaneously. A Pareto optimal solution is defined as 'a set of "non-inferior" solutions in the objective space defining a boundary beyond which none of the objectives can be improved without sacrificing at least one of the other objectives' [7].

An algorithm such as a genetic algorithm, which is commonly used in the literature surrounding Pareto optimisation [15] due to its ability to avoid being trapped in local optima, could be used to search for the Pareto front. This would allow decision-makers to observe the trade-off between the two considered objectives: minimisation of total clinic days scheduled and minimisation of estimated overcollection. Decision-makers at the WBS could then utilise a method, such as the TOPSIS (Technique for Order of Preference by Similarity to Ideal Solution) method [99], to analyse which results are closest to an ideal solution for the WBS. This analysis could inform the most suitable weightings to assign to each term in objective function three of the BDCSM, to ensure that any result would meet their preferences. This method could also be extended to inform decisions for the weights defined as β_1, β_2 and β_3 in the modified objective functions one (4.34) and two (4.35) for the Blood Collection Workforce Scheduling Model (BCWSM).

7.4.2 Consideration of Donor Demographics

As discussed in Section 7.3.5, the consideration of donor demographics can inform better clinic scheduling decisions. The composition of blood types of a given donor panel could be incorporated into the BDCSM or a similar model, alongside blood-type-specific demand figures to better inform the clinic scheduling process. This would provide the potential to better meet demand and to decrease the likelihood of both blood product shortage and expiration. Additionally, analysis of the location of donors and their historical frequency of donation could provide valuable insight to inform strategic decisions such as locations of new clinics for the WBS.

7.4.3 Heuristic Methods

Both the Blood Donation Clinic Scheduling Model and the Blood Collection Workforce Scheduling Model are formulated as a linear progamme and are therefore solved using exact methods. Since the experimental results prove that these models can generate (near) optimal solutions in a reasonable time, there has been no requirement to improve upon the computation times of the models. However, for larger problems such as a longer planning horizon or if the model were to be applied to the blood service of a larger nation (with a greater number of clinics and a higher population), it might be beneficial to consider the use of heuristic methods to enable a good solution to be reached in a reasonable time. Heuristic methods may provide a minimally worse solution but the uncertainty involved in a mobile blood collection process limits how far it can be optimised.

One possible method of expanding the two-stage integer programming model to consider heuristic methods, is to utilise the linear programming relaxation of both the BDCSM and the BCWSM individually as a construction heuristic, similar to various methods presented in the surrounding literature, as discussed in Section 2.2.2. Fischetti et al. [36] discuss various mixed integer linear programming (MILP) based heuristics, including the utilisation of an evolutionary heuristic in combination with a MILP solver, as proposed by Rothberg [83]. Peters et al. [74] present a comparative study between the performance of a MILP and an evolutionary algorithm metaheuristic when used to solve a NP-hard staff assignment problem. The metaheuristic is found to outperform the MILP in only seven minutes. Although the case study considers only 52 staff members, the planning horizon is 400 days, whilst there are over 90 jobs to assign and four different skill levels to consider. A larger instance may take much longer to solve, but is still likely to reach a 'good' solution (within a specified satisfactory optimality gap from the MILP lower-bound) in a reasonable time in comparison to exact methods. This supports the argument that a scaled-up version of both the BDCSM and the BCWSM to consider a larger blood service, with a greater number of both clinics and workers, could be solved effectively utilising a similar evolutionary algorithm metaheuristic.

7.4.4 Intra-Day Scheduling

The literature review in Chapter 2 identified that there is limited research regarding the intra-day scheduling of blood donation clinic staff. Presently at the Welsh Blood Service, all workers assigned to a clinic have a lunch break at the same time, which is enabled by most clinics closing half-way through the working day. This causes a decline in potential productivity with at least an hour of possible donations lost. This suggests an alternative method could be explored regarding intra-day scheduling, where clinics are run at a lower capacity for a period in the day to allow for staggered breaks for the workers.

7.5 Concluding Statement

Throughout this research project, addressing the objectives of the Welsh Blood Service has been a priority, and as a result, a highly specific two-stage integer programming model has been developed to schedule blood donation clinics and the clinic-based workforce more efficiently. This research has addressed gaps in the existing literature surrounding optimisation of the collection echelon in the blood supply chain, such as the tactical scheduling of mobile clinics and clinic-based workers.

Due to the unforeseen impact of the COVID-19 pandemic, unfortunately the model was not able to be implemented at the WBS in the form of a decision support tool, where it would provide the opportunity to minimise costs, overcollection and manual scheduling hours. However, there is still support for implementation of the model at the WBS in the future, including support for utilisation of the model to inform strategic decisions regarding adjustments to the collection model following the pandemic, whilst donors are more accustomed to change.

Bibliography

[1] COIN-OR. https://www.coin-or.org/.

[2] PuLP. https://pypi.org/project/PuLP/.

[3] Nobel Prize Outreach AB 2021. Karl landsteiner - facts. https://www.nobelprize.org/prizes/medicine/1930/landsteiner/facts/>.

[4] A. Ahmadi-Javid, Z. Jalali, and K.J. Klassen. Outpatient appointment systems in healthcare: A review of optimisation studies. *European Journal of Operational Research.*, 258(1):3–34, 2017.

[5] E. Alfonso, V. Augusto, and X. Xie. Mathematical programming models for annual and weekly bloodmobile collection planning. *IEEE Transactions on Automation Science and Engineering*, 12(1):96–105, 2015.

[6] E. Alfonso, X. Xie, V. Augusto, and O. Garraud. Modeling and simulation of blood collection systems. *Health Care Management Science*, 15(1):63–78, 2012.

[7] H.Y. Alhammadi and J.A. Romagnoli. Chapter B4 - Process design and operation: Incorporating environmental, profitability, heat integration and controllability considerations. In P. Seferlis and M.C. Georgiadis, editors, *The Integration of Process Design and Control*, volume 17 of *Computer Aided Chemical Engineering*, pages 264–305. Elsevier, 2004.

[8] M. Arvan, R. Tavakkoli-Moghaddam, and M. Abdollahi. Designing a bi-objective, multi-product supply chain network for blood supply. *Uncertain Supply Chain Management*, 3(1):57–68, 2015.

[9] M.Y.N. Attari, S.H.R. Pasandideh, A. Aghaie, and S.T.A. Niaki. A bi-objective robust optimization model for a blood collection and testing problem: an accelerated stochastic Benders decomposition. *Annals of Operations Research*, (Special Issue: Stochastic Modeling and Optimization, in memory of András Prékopa), 2018.

[10] T. Ayer, C. Zhang, C. Zeng, et al. Analysis and improvement of blood collection operations. *Manufacturing and Service Operations Management*, 21(1):29–46, 2019.

[11] S. Baş Güre, G. Carello, E. Lanzarone, and S. Yalçındağ. Unaddressed problems and research perspectives in scheduling blood collection from donors. *Production Planning and Control*, 29(1):84–90, 2018.

[12] S. Baş, G. Carello, E. Lanzarone, Z. Ocak, and S. Yalçındag. Management of blood donation system: Literature review and research perspectives. *Health Care Systems Engineering for Scientists and Practitioners.*, 169:121–132, 2016.

[13] S. Baş, G. Carello, E. Lanzarone, and S. Yalçındag. An appointment scheduling framework to balance the production of blood units from donation. *European Journal of Operational Research.*, 265(3):1124–1143, 2018.

[14] J. Beliën and H. Forcé. Supply chain management of blood products: A literature review. *European Journal of Operational Research*, 217(1):1–16, 2012.

[15] P. J. Bentley and J. P. Wakefield. Finding acceptable solutions in the pareto-optimal range using multiobjective genetic algorithms. In P. K. Chawdhry, R. Roy, and R. K. Pant, editors, *Soft Computing in Engineering Design and Manufacturing*, pages 231–240. Springer London, 1998.

[16] J. Blake, K. McTaggart, and M. Hardy. Modelling a blood distribution network in the prairies with a generic simulation framework. *Information Systems and Operational Research*, 53(4):194–210, 2015.

[17] J.T. Blake and S. Shimla. Determining staffing requirements for blood donor clinics: the Canadian Blood Services experience. *Transfusion*, 54(3 Pt 2):814–820, 2014.

[18] NHS Blood and Transplant. A history of donation, transfusion and transplantation. `https://www.nhsbt.nhs.uk/who-we-are/a-history-of-donation-transfusion-and-transplantation/`.

[19] NHS Blood and Transplant. How blood is used - blood components. `https://www.blood.co.uk/why-give-blood/how-blood-is-used/blood-components/`.

[20] NHS Blood and Transplant. Our donor centres and local venues. `https://my.blood.co.uk/home`.

[21] S.C. Brailsford, P.R. Harper, B. Patel, and M. Pitt. An analysis of the academic literature on simulation and modelling in health care. *Journal of Simulation*, 3(3):130–140, 2009.

[22] G. Bruno, A. Diglio, C. Piccolo, and L. Cannavacciuolo. Territorial reorganization of regional blood management systems: Evidences from an Italian case study. *Omega (United Kingdom)*, 89:54–70, 2019.

[23] E. Cetin and L.S. Sarul. Blood bank location model: A multiobjective approach. *European Journal of Pure and Applied Mathematics*, 2(1):112–124, 2009.

[24] P. Chaiwuttisak, H. Smith, Y. Wu, et al. Location of low-cost blood collection and distribution centres in Thailand. *Operations Research for Health Care*, 9:7–15, 2016.

[25] S. Cheraghi and S.-M. Hosseini-Motlagh. Responsive and reliable injured-oriented blood supply chain for disaster relief: a real case study. *Annals of Operations Research*, 291:129–167, 2020.

[26] American Red Cross. History of blood transfusion. `https://www.redcrossblood.org/donate-blood/blood-donation-process/what-happens-to-donated-blood/blood-transfusions/history-blood-transfusion.html`.

[27] B. Custer, E.S. Johnson, S.D. Sullivan, et al. Community blood supply model: Development of a new model to assess the safety, sufficiency, and cost of the blood supply. *Medical Decision Making*, 25(5):571–582, 2005.

[28] L. Dean. *Blood Groups and Red Cell Antigens.* National Center for Biotechnology Information (US), 2005.

[29] R.E. Domen, I.D. Grewal, N.V. Hirschler, and G.A. Hoeltge. An evaluation of the need for shared blood donor deferral registries. *International Journal for Quality in Health Care*, 9(1):35–41, 1997.

[30] M. Doneda, S. Yalçındağ, I. Marques, and E. Lanzarone. A discrete-event simulation model for analysing and improving operations in a blood donation centre. *Vox Sanguinis*, 116(10):1060–1075, 2021.

[31] H. Ellis. James Blundell, pioneer of blood transfusion. *British Journal of Hospital Medicine*, 68(8):477, 2007.

[32] S. Eom and E. Kim. A survey of decision support system applications (1995–2001). *Journal of the Operational Research Society*, 57(11):1264–1278, 2006.

[33] B. Fahimnia, A. Jabbarzadeh, A. Ghavamifar, and M. Bell. Supply chain design for efficient and effective blood supply in disasters. *International Journal of Production Economics*, 183:700–709, 2017.

[34] E. Farrokhizadeh, S.A. Seyfi-Shishavan, and S.I. Satoglu. Blood supply planning during natural disasters under uncertainty: a novel bi-objective model and an application for Red Crescent. *Annals of Operations Research*, (Special Issue: Design and Management of Humanitarian Supply Chains), 2021.

[35] M. Fazli-Khalaf, S. Khalilpourazari, and M. Mohammadi. Mixed robust possibilistic flexible chance constraint optimization model for emergency blood supply chain network design. *Annals of Operations Research*, 283:1079–1109, 2019.

[36] M. Fischetti and A. Lodi. *Heuristics in Mixed Integer Programming.* John Wiley and Sons, Ltd, 2011. Wiley Encyclopedia of Operations Research and Management Science.

[37] Michigan Center for Healthcare Engineering and Patient Safety. Shift scheduling game. `https://cheps.engin.umich.edu/tools/shift-scheduling-game/`.

[38] J.L. Gearhart, K.L. Adair, J.D. Durfee, et al. Comparison of open-source linear programming solvers. U.S. Department of Energy; Office of Scientific and Technical Information, 2013.

[39] M.R. Ghatreh Samani, S.A. Torabi, and S.-M. Hosseini-Motlagh. Integrated blood supply chain planning for disaster relief. *International Journal of Disaster Risk Reduction*, 27:168–188, 2018.

[40] Welsh Government. Summary statistics for Wales, by region: 2020. `https://gov.wales/sites/default/files/statistics-and-research/2020-05/summary-statistics-regions-wales-2020-629.pdf`, 2020.

[41] S. Gunpinar and G. Centeno. An integer programming approach to the bloodmobile routing problem. *Transportation Research Part E: Logistics and Transportation Review*, 86:94–115, 2016.

[42] M. Habibi-Kouchaksaraei, M.M. Paydar, and E. Asadi-Gangraj. Designing a bi-objective multi-echelon robust blood supply chain in a disaster. *Applied Mathematical Modelling*, 55:583–599, 2018.

[43] A. Haeri, S.-M. Hosseini-Motlagh, M.R. Ghatreh Samani, and M. Rezaei. A mixed resilient-efficient approach toward blood supply chain network design. *International Transactions in Operational Research*, 27(4):1962–2001, 2020.

[44] B. Hamdan and A. Diabat. A two-stage multi-echelon stochastic blood supply chain problem. *Computers and Operations Research*, 101:130–143, 2019.

[45] P.R. Harper and M.A. Pitt. On the challenges of healthcare modelling and a proposed project life cycle for successful implementation. *Journal of the Operational Research Society*, 55(6):657–661, 2004.

[46] S.-M. Hosseini-Motlagh, M.R.G. Samani, and S. Cheraghi. Robust and stable flexible blood supply chain network design under motivational initiatives. *Socio-Economic Planning Sciences*, 70:100725, 2020.

[47] S.-M. Hosseini-Motlagh, M.R.G. Samani, and S. Homaei. Blood supply chain management: robust optimization, disruption risk, and blood group compatibility (a real-life case). *Journal of Ambient Intelligence and Humanized Computing*, 11(3):1085–1104, 2020.

[48] C.L. Hsieh. An evolutionary-based optimization for a multi-objective blood banking supply chain model. *Modern Advances in Applied Intelligence*, 8481:511–520, 2014.

[49] P. Hulshof, N. Kortbeek, R. Boucherie, et al. Taxonomic classification of planning decisions in health care: a structured review of the state of the art in OR/MS. *Health Systems*, 1(2):129–175, 2012.

[50] D.A. Jacobs, M.N. Silan, and B.A. Clemson. An analysis of alternative locations and service areas of american red cross blood facilities. *Interfaces*, 26(3):40–50, 1996.

[51] S. Khalilpourazari and A. Arshadi Khamseh. Bi-objective emergency blood supply chain network design in earthquake considering earthquake magnitude:

a comprehensive study with real world application. *Annals of Operations Research*, 283:355–393, 2019.

[52] S. Khalilpourazari, S. Soltanzadeh, G.-W. Weber, and S.K. Roy. Designing an efficient blood supply chain network in crisis: neural learning, optimization and case study. *Annals of Operations Research*, 289(1):123–152, 2020.

[53] R. Kolisch and A. Drexl. Local search for nonpreemptive multi-mode resource-constrained project scheduling. *IIE Transactions*, 29:987–999, 1997.

[54] O. Koné, C. Artigues, P. Lopez, and M. Mongeau. Event-based MILP models for resource-constrained project scheduling problems. *Computers and Operations Research*, 38(1):3–13, 2011.

[55] A. Kushniruk. Evaluation in the design of health information systems: application of approaches emerging from usability engineering. *Computers in Biology and Medicine*, 32(3):141–149, 2002.

[56] W.C. Lee and B.W. Cheng. An intelligent system for improving performance of blood donation. *Journal of Quality*, 18(2):173–185, 2011.

[57] R. Lewisohn. The importance of the proper dosage of sodium citrate in blood transfusion. *Annals of Surgery*, 64(5):618–623, 1916.

[58] X. Li, H. Fan, J. Liu, and Q. Xun. Staff scheduling in blood collection problems. *Annals of Operations Research*, (Special Issue: Scalable Optimization and Decision Making in OR), 2020.

[59] E.J. Lodree, D. Carter, and E. Barbee. The donation collections routing problem. *Springer Proceedings in Mathematics and Statistics*, 185:159–189, 2016.

[60] H. Lowalekar, R. Nilakantan, and N. Ravichandran. Analysis of an order-up-to-level policy for perishables with random issuing. *Journal of the Operational Research Society*, 67(3):483–505, 2016.

[61] H. Lowalekar and R.R. Ravi. Revolutionizing blood bank inventory management using the TOC thinking process: An Indian case study. *International Journal of Production Economics*, 186:89–122, 2017.

[62] H. Lowalekar and N. Ravichandran. Model for blood collections management. *Transfusion*, 50(12 Pt 2):2778–2784, 2010.

[63] A. Mobasher, A. Ekici, and O.Ö. Özener. Coordinating collection and appointment scheduling operations at the blood donation sites. *Computers and Industrial Engineering*, 87:260–266, 2015.

[64] M. Mugdh and S. Pilla. A conceptual framework for achieving balance between innovation and resilience in optimizing emergency department operations. *The Health Care Manager*, 30(4):352–360, 2011.

[65] A. Nagurney and P. Dutta. Competition for blood donations. *Omega (United Kingdom)*, 85:103–114, 2019.

[66] NHS. NHS agenda for change 2020/21 pay scales. https://www.healthcareers.nhs.uk/working-health/working-nhs/nhs-pay-and-benefits/agenda-change-pay-rates/.

[67] World Health Organization. Blood safety and availability. https://www.who.int/news-room/fact-sheets/detail/blood-safety-and-availability, 2020.

[68] A.F. Osorio, S.C. Brailsford, and H.K. Smith. A structured review of quantitative models in the blood supply chain: A taxonomic framework for decision-making. *International Journal of Production Research*, 53(24):7191–7212, 2015.

[69] A.F. Osorio, S.C. Brailsford, and H.K. Smith. Whole blood or apheresis donations? A multi-objective stochastic optimization approach. *European Journal of Operational Research*, 266(1):193–204, 2018.

[70] A.F. Osorio, S.C. Brailsford, H.K. Smith, and J. Blake. Designing the blood supply chain: how much, how and where? *Vox Sanguinis*, 113(8):760–769, 2018.

[71] A.F. Osorio, S.C. Brailsford, H.K. Smith, S.P. Forero-Matiz, and B.A. Camacho-Rodríguez. Simulation-optimization model for production planning in the blood supply chain. *Health Care Management Science*, 20(4):548–564, 2017.

[72] O. Özener, A. Ekici, and E. Çoban. Improving blood products supply through donation tailoring. *Computers and Operations Research*, 102:10–21, 2019.

[73] O.Ö. Özener and A. Ekici. Managing platelet supply through improved routing of blood collection vehicles. *Computers and Operations Research*, 98:113–126, 2018.

[74] J. Peters, D. Stephan, I. Amon, H. Gawendowicz, et al. Mixed integer programming versus evolutionary computation for optimizing a hard real-world staff assignment problem. *Proceedings of the International Conference on Automated Planning and Scheduling*, 29(1):541–554, 2021.

[75] A. Pirabán, W.J. Guerrero, and N. Labadie. Survey on blood supply chain management: Models and methods. *Computers and Operations Research*, 112:104756, 2019.

[76] M. Pouraliakbari-Mamaghani, A. Ghodratnama, S.H.R. Pasandideh, and A. Saif. A robust possibilistic programming approach for blood supply chain network design in disaster relief considering congestion. *Operational Research - An International Journal*, 2021. Advance online publication.

[77] M. Rabbani, M. Aghabegloo, and H. Farrokhi-Asl. Solving a bi-objective mathematical programming model for bloodmobiles location routing problem. *International Journal of Industrial Engineering Computations*, 8(1):19–32, 2017.

[78] D. Rahmani. Designing a robust and dynamic network for the emergency blood supply chain with the risk of disruptions. *Annals of Operations Research*, 283(1-2):613–641, 2019.

[79] R. Ramezanian and Z. Behboodi. Blood supply chain network design under uncertainties in supply and demand considering social aspects. *Transportation Research Part E: Logistics and Transportation Review*, 104:69–82, 2017.

[80] D. Ranum, B. Miller, J. Zelle, and M. Guzdial. Successful approaches to teaching introductory computer science courses with Python. In *Proceedings of the 37th SIGCSE Technical Symposium on Computer Science Education*, SIGCSE '06, page 396–397, New York, NY, USA, 2006. Association for Computing Machinery.

[81] N. Razavi, H. Gholizadeh, S. Nayeria, and T.A. Ashrafi. A robust optimization model of the field hospitals in the sustainable blood supply chain in crisis logistics. *Journal of the Operational Research Society*, 72(12):2804–2828, 2021.

[82] N. Roberts, S. James, M. Delaney, and C. Fitzmaurice. The global need and availability of blood products: a modelling study. *The Lancet Haematology*, 6(12):E606–E615, 2019.

[83] E. Rothberg. An evolutionary algorithm for polishing mixed integer programming solutions. *INFORMS Journal on Computing*, 19(4):534–541.

[84] G. Sahin, H. Süral, and S. Meral. Locational analysis for regionalization of Turkish Red Crescent blood services. *Computers and Operations Research.*, 34(3):692–704, 2007.

[85] F. G. Şahinyazan, B.Y. Kara, and M.R. Taner. Selective vehicle routing for a mobile blood donation system. *European Journal of Operational Research*, 245(1):22–34, 2015.

[86] F. Salehi, M. Mahootchi, and S.M.M. Husseini. Developing a robust stochastic model for designing a blood supply chain network in a crisis: a possible earthquake in Tehran. *Annals of Operations Research*, 283:679–703, 2019.

[87] M.R.G. Samani and S.-M. Hosseini-Motlagh. A novel capacity sharing mechanism to collaborative activities in the blood collection process during the COVID-19 outbreak. *Applied Soft Computing*, 112:107821, 2021.

[88] M.R.G. Samani, S.-M. Hosseini-Motlagh, and S. Homaei. A reactive phase against disruptions for designing a proactive platelet supply network. *Transportation Research Part E: Logistics and Transportation Review*, 140:102008, 2020.

[89] M.R.G. Samani and S.M. Hosseini-Motlagh. A robust framework for designing blood network in disaster relief: a real-life case. *Operational Research - An International Journal*, 21(3):1529–1568, 2021.

[90] M.R.G. Samani, S.M. Hosseini-Motlagh, and S.F. Ghannadpour. A multilateral perspective towards blood network design in an uncertain environment: Methodology and implementation. *Computers and Industrial Engineering*, 130:450–471, 2019.

[91] S.A. Seyfi-Shishavan, Y. Donyatalab, E. Farrokhizadeh, and S.I. Satoglu. A fuzzy optimization model for designing an efficient blood supply chain network under uncertainty and disruption. *Annals of Operations Research*, 2021. Advance online publication.

[92] L. Shi, J. Wang, L. Stevens, P. Ness, and H. Shan. Blood safety and availability: continuing challenges in China's blood banking system. *Transfusion*, 54(2):471–482, 2014.

[93] M. Soltani, R.A. Mohammadi, S.M.H. Hosseini, and M.M. Zanjani. A new model for blood supply chain network design in disasters based on hub location approach considering intercity transportation. *International Journal of Industrial Engineering and Production Research*, 32(2):1–28, 2021.

[94] O. Tange. GNU Parallel 2018. `https://doi.org/10.5281/zenodo.1146014`, 2018.

[95] R. Tavakkoli-Moghaddam, S. Mortazavi, S.H. Madani-Saatchi, and M. Fallah. Designing a bi-objective model for a blood bank location-allocation problem in emergency logistics under uncertainty. 2016. CIIE 2016: 46th Iternational Conference on Computers and Industrial Engineering.

[96] M.C. Testik, B.Y. Ozkaya, S. Aksu, and O.I. Ozcebe. Discovering blood donor arrival patterns using data mining: A method to investigate service quality at blood centers. *Journal of Medical Systems*, 36(2):579–594, 2012.

[97] Theory of Constraints Institute. `https://www.tocinstitute.org/`.

[98] S.P.J. van Brummelen. Bloody fast blood collection, 2017. PhD thesis. University of Twente.

[99] Z. Wang and G.P. Rangaiah. Application and analysis of methods for selecting an optimal solution from the Pareto-optimal front obtained by multiobjective optimization. *Industrial and Engineering Chemistry Research*, 56(2):560–574.

[100] J. Webster and T.J. Watson. Analyzing the past to prepare for the future: Writing a literature review. *Management Information Systems Quarterly*, 26(2):13–23, 2002.

[101] E.P. Williams, P.R. Harper, and D. Gartner. Modeling of the collections process in the blood supply chain: A literature review. *IISE Transactions on Healthcare Systems Engineering*, 10(3):200–211, 2020.

[102] Y. Xu, A.M. Wood, M.J. Sweeting, D.J. Roberts, and B.D.M. Tom. Optimal individualized decision rules from a multi-arm trial: A comparison of methods and an application to tailoring inter-donation intervals among blood donors in the UK. *Statistical Methods in Medical Research*, 29(11):3113–3134, 2020.

[103] S. Yalçındağ, S.B. Güre, G. Carello, and E. Lanzarone. A stochastic risk-averse framework for blood donation appointment scheduling under uncertain donor arrivals. *Health Care Management Science*, 23(4):535–555, 2020.

[104] B. Zahiri and M.S. Pishvaee. Blood supply chain network design considering blood group compatibility under uncertainty. *International Journal of Production Research*, 55(7):2013–2033, 2017.

[105] B. Zahiri, S.A. Torabi, M. Mousazadeh, and S.A. Mansouri. Blood collection management: Methodology and application. *Applied Mathematical Modelling*, 39(23-24):7680–7696, 2013.

Appendix A

Literature Review Search String

TITLE-ABS-KEY("blood service" OR "blood donation" OR "blood collection" OR "blood type*" OR "blood clinic" OR "donation clinic" OR "blood donor" OR "blood donor appointment scheduling" OR "blood service appointment scheduling" OR "blood clinic appointment scheduling" OR "blood donation appointment scheduling") ("clustering" OR "k-means" OR "Attribute selection" OR "classification" OR "feature selection" OR "machine learning" OR "multidimensional" OR "Bayes*" OR "combined classification" OR "correlation-based" OR "decision rule" OR "information gain" OR "markov blanket" OR "neural network" OR "principal component" OR "relief" OR "support vector machine" OR "tree" OR "wrapper" OR "optimization" OR "optimisation" OR "scheduling" OR "mixed integer program*" OR "simulation" OR "heuristics" OR "L-shaped" OR "programming" OR "mathematical program*" OR "integer program*" OR "column generation" OR "constraint program*" OR "linear program*" OR "branch and bound" OR "branch and price" OR "dynamic program*" OR "goal program*" OR "quadratic program*") AND ISSN (0001-5172 OR 0003-2409 OR 0310-057x OR 0003-2999 OR 0003-3022 OR 0007-0912 OR 0832-610x OR 0749-8047 OR 0952-7907 OR 0265-0215 OR 1090-3801 OR 0959-289x OR 0913-8668 OR 1053-0770 OR 0952-8180 OR 1387-1307 OR 0898-4921 OR 0304-3959 OR 1530-7085 OR 1155-5645 OR 1098-7339

OR 0003-6870 OR 0007-8506 OR 0360-8352 OR 0305-0548 OR 0013-791x OR 0014-0139 OR 1751-5254 OR 1077-2618 OR 0018-9391 OR 1551-3203 OR 0740-817x OR 1542-894x OR 0263-5577 OR 0143-991x OR 1943-670x OR 0169-8141 OR 0925-5273 OR 0020-7543 OR 0748-5492 OR 0733-9364 OR 0923-4748 OR 0742-597x OR 0278-6125 OR 0924-0136 OR 0737-6782 OR 0022-4065 OR 0269-9648 OR 1748-006x OR 0953-7287 OR 0898-2112 OR 0748-8017 OR 1684-3703 OR 0746-9179 OR 0951-8320 OR 0934-9839 OR 0895-6308 OR 0925-7535 OR 1012-277x OR 1098-1241 OR 0166-4972 OR 0933-3657 OR 1472-6947 OR 1538-2931 OR 0169-2607 OR 1460-4582 OR 1833-3583 OR 0739-5175 OR 1089-7771 OR 1753-8157 OR 1386-5056 OR 0266-4623 OR 1532-0464 OR 0885-8195 OR 1356-1294 OR 1438-8871 OR 0148-5598 OR 1067-5027 OR 0140-0118 OR 0272-989x OR 0026-1270 OR 0962-2802 OR 0277-6715 OR 1619-4500 OR 0254-5330 OR 1524-1904 OR 0217-5959 OR 1435-246x OR 0926-6003 OR 0305-0548 OR 1063-293x OR 0167-9236 OR 0924-6703 OR 1572-5286 OR 0305-215x OR 1751-5254 OR 0377-2217 OR 0957-4174 OR 1936-6582 OR 1568-4539 OR 1932-8184 OR 0740-817x OR 0315-5986 OR 1091-9856 OR 0092-2102 OR 0951-192x OR 0219-6220 OR 0925-5273 OR 0020-7543 OR 0020-7721 OR 0267-5730 OR 0969-6016 OR 1004-3756 OR 0925-5001 OR 1547-5816 OR 0278-6125 OR 0272-6963 OR 0022-3239 OR 0022-4065 OR 1094-6136 OR 1004-4132 OR 0160-5682 OR 0453-4514 OR 0025-1909 OR 1523-4614 OR 1432-2994 OR 0025-5610 OR 0364-765x OR 0275-5823 OR 0894-069x OR 0028-3045 OR 1566-113x OR 0305-0483 OR 0030-364x OR 0167-6377 OR 0233-1934 OR 0143-2087 OR 1389-4420 OR 1862-4472 OR 1055-6788 OR 0171-6468 OR 1348-9151 OR 0269-9648 OR 1059-1478 OR 0953-7287 OR 0748-8017 OR 0257-0130 OR 0399-0559 OR 0951-8320 OR 0925-7535 OR 1052-6234 OR 1696-2281 OR 1220-1766 OR 0167-6911 OR 1098-1241 OR 0166-4972 OR 1134-5764 OR 0191-2615 OR 1366-5545 OR 0041-1655 OR 1748-006x OR 0013-791x OR 0894-587x OR 0954-0121 OR 1088-0224 OR 1819-5164 OR 1448-7527 OR 1472-698x OR 2044-5415 OR 0963-1801 OR 0010-3853 OR 1936-6574 OR 1618-7598 OR 0163-2787 OR 1054-8289 OR 0278-2715 OR 1065-3058 OR 0195-8631 OR 1386-9620 OR 0361-6274 OR 1041-0236 OR 1057-9230 OR 1744-1331 OR 1369-6513 OR 1833-3583 OR

0168-8510 OR 0268-1080 OR 0957-4824 OR 1477-7525 OR 0017-9124 OR 1446-1242 OR 1478-4491 OR 1748-5908 OR 0046-9580 OR 1389-6563 OR 0749-6753 OR 0020-7314 OR 1353-4505 OR 0898-2643 OR 1094-3412 OR 0094-5145 OR 1049-2089 OR 0167-6296 OR 0361-6878 OR 1355-8196 OR 1096-9012 OR 1356-1820 OR 1091-4358 OR 0825-8597 OR 0891-5245 OR 1741-1122 OR 0197-5897 OR 0890-765x OR 0025-7079 OR 1077-5587 OR 0887-378x OR 1178-1653 OR 1170-7690 OR 1076-8971 OR 1075-2730 OR 1049-7323 OR 1729-0376 OR 1098-3015)

Appendix B

Actual Clinic Schedule Figures

Table B.1: Actual Number of Clinics and Overcollection over the Planning Horizons

	Planning Horizon 1 07/01/2019 - 03/02/2019		**Planning Horizon 2 22/07/2019 - 18/08/2019**	
Region	**Number of Clinics Scheduled**	**Estimated Overcollection**	**Number of Clinics Scheduled**	**Estimated Overcollection**
AW	119	1080	109	63
NE	16	*160*	15	-*22*
NW	16	*228*	15	-*3*
SE	72	*498*	64	*64*
SW	15	*194*	15	*23*

Appendix C

Blood Collection Workforce Scheduling Model: Original vs. Modified Model

Figure C.1: BCWSM Computational Times - Original vs. Modified Model

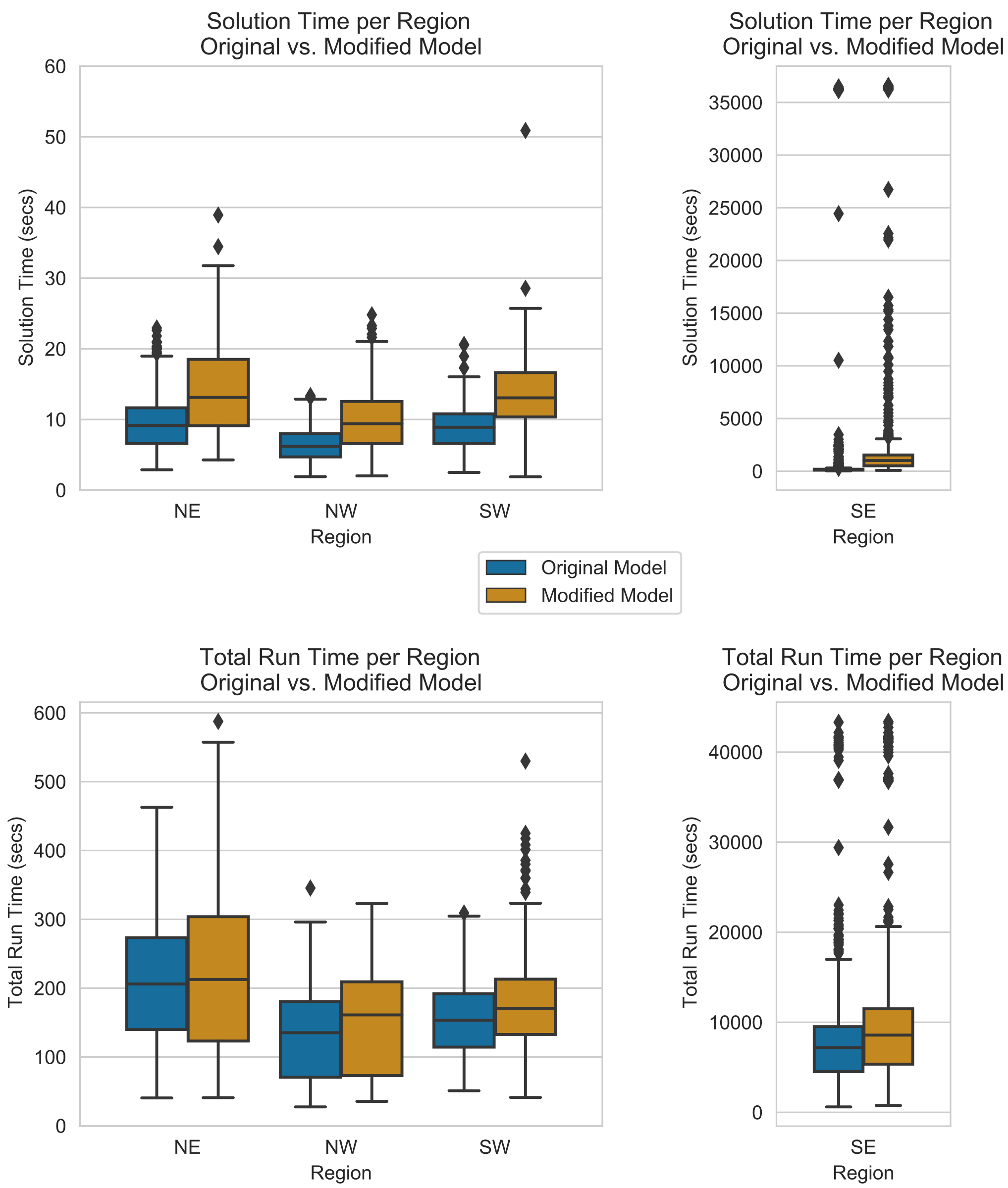
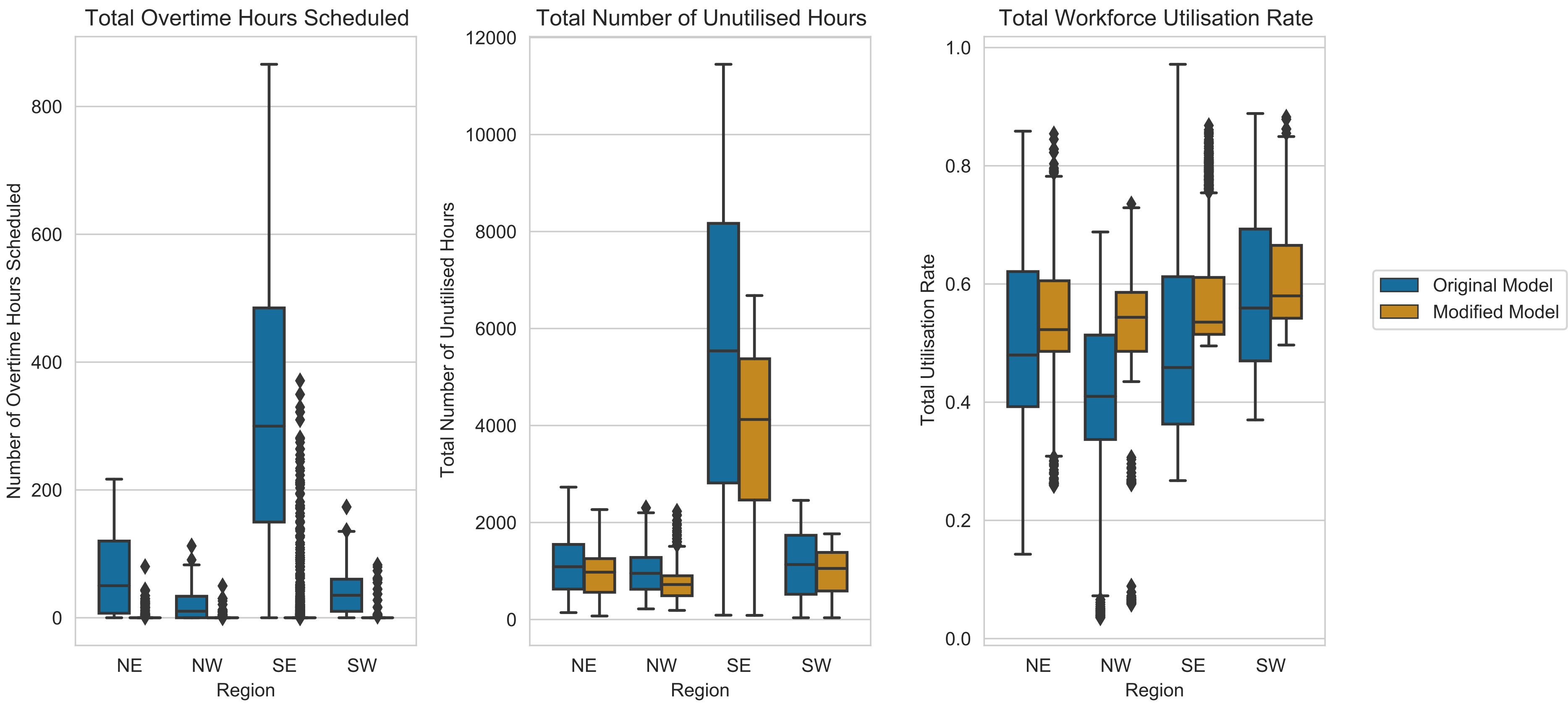


Figure C.2: BCWSM Solutions - Original vs. Modified Model



Appendix D

Clinic Duration Matrix Examples

Figure D.1: Clinic Duration Matrix - Consecutive Day Clinics

```
###########################################################
# clinic duration matrices

# for one-day clinics
clinic_1_day = np.identity(number_of_days, dtype=int)

# for clinics on two consecutive days
clinic_2_day = np.eye(number_of_days, dtype=int)

for starting_day in range(number_of_days - 1):
    clinic_2_day[starting_day][starting_day + 1] = 1
    clinic_2_day[-1] = np.zeros(number_of_days)

# for clinics on three consecutive days
clinic_3_day = np.eye(number_of_days, dtype=int)

for starting_day in range(number_of_days - 2):
    clinic_3_day[starting_day][starting_day + 1] = 1
    clinic_3_day[starting_day][starting_day + 2] = 1
    clinic_3_day[-1] = np.zeros(number_of_days)
    clinic_3_day[-2] = np.zeros(number_of_days)

# for clinics on four consecutive days
clinic_4_day = np.eye(number_of_days, dtype=int)

for starting_day in range(number_of_days - 3):
    clinic_4_day[starting_day][starting_day + 1] = 1
    clinic_4_day[starting_day][starting_day + 2] = 1
    clinic_4_day[starting_day][starting_day + 3] = 1
    clinic_4_day[-1] = np.zeros(number_of_days)
    clinic_4_day[-2] = np.zeros(number_of_days)
    clinic_4_day[-3] = np.zeros(number_of_days)
```

Figure D.2: Clinic Duration Matrix - One Clinic for Two Consecutive Weeks

```
# for clinics that occur for one day a week for two consecutive weeks
clinic_1_per_week_2 = np.array(
    [
        [1,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0],
        [0,1,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0],
        [0,0,1,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0],
        [0,0,0,1,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0],
        [0,0,0,0,1,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0],
        [0,0,0,0,0,1,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0],
        [0,0,0,0,0,0,1,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0],
        [0,0,0,0,0,0,0,1,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0],
        [0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0],
        [0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0],
        [0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0],
        [0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0],
        [0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0],
        [0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,1,0,0,0,0,0,0,0],
        [0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,1,0,0,0,0,0,0],
        [0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,1,0,0,0,0,0],
        [0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,1,0,0,0,0],
        [0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,1,0,0,0],
        [0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,1,0,0],
        [0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,1,0],
        [0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,1],
        [0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0],
        [0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0],
        [0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0],
        [0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0],
        [0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0],
        [0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0],
        [0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0]
    ]
```