Time and Timing in Vehicle Routing Problems

This thesis is number D138 of the thesis series of the Beta Research School for Operations Management and Logistics. The Beta Research School is a joint effort of the departments of Industrial Engineering & Innovation Sciences, and Mathematics and Computer Science at Eindhoven University of Technology and the Centre for Production, Logistics and Operations Management at the University of Twente.

A catalogue record is available from the Eindhoven University of Technology Library.

ISBN: 978-90-386-2348-1

Printed by University Printing Office, Eindhoven

This research has been funded by the Netherlands Organization for Scientific Research (NWO), project number 400-05-185.

Time and Timing in Vehicle Routing Problems

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de
Technische Universiteit Eindhoven, op gezag van de
rector magnificus, prof.dr.ir. C.J. van Duijn, voor een
commissie aangewezen door het College voor
Promoties in het openbaar te verdedigen
op woensdag 24 november 2010 om 16.00 uur

door

Ola Jabali

geboren te Nazareth, Israël

Dit proefschrift is goedgekeurd door de promotor:

prof.dr. A.G. de Kok

Copromotor:
dr. T. van Woensel

Aileme...
özellikle,
yeğenime

To my family...
especially,
to my nephew

Acknowledgements

There are two major manners of telling an event or situation: Through a short story or a long story. My friends know... well, not only know but also complain about the fact that I tell long stories. If you ask something... anything... my response can start with something like the big bang theory. I will also tell a story regarding this thesis now. I will start the story with some parts from its end, just in case you are so busy (bored) and cannot read the full story.

I guess that all Ph.D. students get frustrated every now and then, as they often get lost in the chaotic clouds of thoughts and ideas. My first promoter, Geert-Jan van Houtum, has been a great guide, clarifying the paths that I could follow whenever I was lost. I have never had a question left without an answer. He helped me zoom in and out on problems. He also taught me a lot about how to write down my work properly. Thank you so so much Geert-Jan.

It has been a great fortune to me to have Gudrun Kiesmüller as my daily supervisor. Her office was just a few meters away from mine and she was so open to any discussion such that I often had the chance to knock on her door and start talking without an appointment. Her perspective, which was different from Geert-Jan's, improved my approach towards the problems considerably. I enjoyed the time that I spent with her in front of the board for formulations and proofs. Gudrun: Thank you for all your effort and friendly supervision.

I had a visit to Carnegie Mellon University (CMU) in Pittsburgh, USA, between March-June 2009. I worked under the supervision of Alan Scheller-Wolf. Again, a different perspective which was exhibited in a very positive manner. Alan had a very busy agenda... and I was always in that agenda. The third chapter of this thesis includes the research that I started at CMU. My stay there was an extraordinary experience, not only in terms of research. Things were different in the USA... different from what I was thinking of. Thank you Alan: First, for your supervision and cooperation; second, for providing me the opportunity to get rid of some of my prejudices and establish very good friendships there.

Now, back to the beginning...

I remember a conversation between me and my mom, some time after I started the first grade. She asked me how I was feeling about the school (apparently, after observing me for a while). I did not have any troubles with the classes. I was not crying after my mom or something. But I did not like the school. It was so boring. I told her that I still wanted to be a free kid, playing all the time, drawing pictures, etc, etc, etc. The school education started too early for me (I was 6)! The next day, I was a free kid. I thank to my parents, because not only they are my parents and take care of me (I am almost 33 but they are still on duty), but also they have always respected my opinions and decisions.

-Canım annem ve babam, sizlere çok çok teşekkür ediyorum. Sadece annem-babam olmanızdan dolayı değil, aynı zamanda küçücük çocukken dahi düşünce ve kararlarıma saygı göstermenizden dolayı.-

I restarted the primary school next year. This time, I was ready for it. I loved math from the beginning. I still love it. If you have a look at the further pages of this book, you will see some math. I am grateful to my primary school teacher, Ahmet Hüseyin Petek, for reinforcing my interests. I know quite many people who hated most of the things taught in classrooms. And I also know that a large portion of their hate stemmed from the manner that they were taught.

I also feel grateful to (almost) all my high school teachers, but I would like to mention two of them in particular. I did not get the best scores in English classes, but thanks to my English teacher, Serpil Mıstık, I can write a book fully in English now. I and many friends of mine have a great gratitude to our math teacher, Mustafa Özdemir, who made us study hard for the university entrance exam at an early stage compared to other students... and we all made it to very good universities.

I had a great time at Boğaziçi University. I was studying hard. I was also enjoying the life in the beautiful campus on the coast of Bosphorus. Thanks to all academic members at the Industrial Engineering Department of Boğaziçi University. Special thanks to my supervisor there, Kuban Altınel, who has a very large impact on my views on science. I will never forget his support which was not limited only to my academic life.

The education at the Industrial Engineering Department of Boğaziçi University provides a very strong theoretical background. However, I think that there is a lack of bridging theory and practice in the education (indeed, I think that this is a general problem in the education system in Turkey). This connection has been established throughout my Ph.D. research. Thanks to the Innovation-Oriented Research Programme 'Integrated Product Creation and Realization (IOP IPCR)' of the Netherlands Ministry of Economic Affairs, which funded my research and enabled me to cooperate with a number of companies. I would also like to thank to Ad Zephat and Guillaume Stollman from Philips Healthcare; and Wout Smans and Radj Bachoe from Vanderlande Industries. I can confidently state the motivation of the problems that I tackled in this thesis and the use of our research by their help.

I would like to thank my second promoter, Ton de Kok, for his high-level comments and ideas about my research. Thanks to Martin Newby for his valuable comments about my thesis. His remarks made me really think of the problems in a different way.

At the department, I sometimes walked around in the corridor annoyingly... distracting my colleagues (friends!). Thanks to all for their tolerance. But, unfortunately, no-one could save Michiel Jansen, my office-mate, from suffering my existence all the time. So, you understand that he deserves to be thanked separately. I would also like to thank Ingrid van Helvoort - Vliegen for being my 'supervisor' about the Dutch way of life and education system, especially in my first year here. She also helped me with the writing of this thesis.

There are three other friends that I have to mention here. Especially two of them would kill me if I do not do so. And, honestly, they have such a right... to some extent. First: Thank you Kostas Kevrekidis for our discussions and cooperation... and, of course, your great friendship. Then: Thank you Ola Jabali (Olla Gabali is a misspelling!) and Hayriye Çağnan, my angels who saved my life multiple times, one of which was related to the numerical precision problem that I had with the fourth chapter of this thesis. Now... you know who have... some right.

I told you, I tell long stories. But, if these people were not involved in my life and these events did not happen, you couldn't read this book, if you ever intended to do so (in the meanwhile, there were other steps in between which I also appreciate). I do not mean that there would not be any book written by me. But, it could not be this one. I have been a lucky man so far. And I hope that I will be lucky enough to tell some stories in the future.

At the very end, I would also like to thank Matthieu van der Heijden and Fred Langerak for participating in my Ph.D. committee.

Contents

Chapter 1

Introduction

The pages this thesis is printed on have been transported from a distribution center in Zutphen to the Eindhoven University of Technology (TU/e). The distance between these two locations is 118 km. Consider that the warehouse in Zutphen has a number of vehicles used to transport office material to a number of locations. Choosing *which* of these vehicles to dispatch to the TU/e and determining *when* this vehicle should arrive at the TU/e, is precisely the topic of this thesis.

In 1997, 3.7 billion tons where transported on the roads of the USA, nearly the double of the number in 1984 (**?**). In 2006, road freight transport was about 1894 billion tonne-kilometres in the European Union. Furthermore, in Europe the total road freight transport volume increased by 43% between 1992 and 2005 (**?**). One of the main drives of this trend is the increasing demand for goods transport.

In the context of road freight transport, this thesis considers short-haul transportation. This concerns the pick-up and delivery of goods in a small area, i.e, a city or a country (**?**). In general, vehicles are based at a single depot. A vehicle conducts one tour per working day, performing pick or delivery services. In short-haul transportation, the key strategic question relates to the location of the depot. On a tactical level decisions concerning fleet mix and size need to be taken. Finally, at the operational level lies the Vehicle Routing Problem (VRP). The VRP addresses the issue of constructing routes, for a given fleet, to satisfy customer requirements.

The VRP has a high relevance to real-life practice. Thus, much effort is invested into developing efficient solution procedures to the problem. The basic formulation of the VRP does not cover all complexities encountered in real-life. This thesis strives to introduces several realistic variants to the VRP. The proposed variants mainly elaborate on the time it takes a vehicle to complete a route and on its expected arrival time at customers.

The rest of this chapter is organized as follows. Section **??** defines the Vehicle Routing Problem. Section **??** discusses the applications that motivated the research conducted in the thesis. Section **??** presents the uncertainty aspects tackled in the thesis. Section **??** presents an overview of the main chapters. Section **??** summarizes the solution methods used in the thesis. Finally, in Section **??** the outline of the rest of the thesis is given.

1.1. The Vehicle Routing Problem

The Vehicle Routing Problem (VRP), as first introduced by **?**, involves constructing optimal delivery routes from a central depot to a set of geographically dispersed customers. Its wide applicability and inherent complexity lead to an extensive amount of research. In what follows, we formally define the VRP (see also **?** for a general survey).

The VRP, sometimes referred to as capacitated VRP, is defined on a complete undirected graph $G = (V, A)$ where $V = \{0, 1, \ldots, n\}$ is a set of vertices and $A = \{(i, j) : i <> j \in V\}$ is the set of arcs. The vertex 0 denotes the depot; the other vertices in V represent customers. The problem considers using m identical vehicles each with a fixed capacity Q. Each customer has a non-negative demand $q_i \leq Q$. Associated with each edge $(v_i, v_j) \in A$, c_{ij} denotes the cost of visiting node j immediately after visiting node i. The costs are assumed to be symmetric, i.e., $c_{ij} = c_{ji}$ for all i,j. The objective is to design routes for the m vehicles yielding minimum cost where the following conditions hold:

1. Each customer is visited exactly once by exactly one vehicle.
2. All vehicle routes start and end at the single depot.
3. Every route has a total demand not exceeding the vehicle capacity Q.

Several mathematical models were proposed to describe the VRP. We present the two-index vehicle flow formulation based on the work of **?**. Let x_{ij} be an integer variable representing the number of times edge $[i, j]$ appears in the solution. x_{ij} is binary for $i, j \in V \setminus \{0\}$. For all $j \in V \setminus \{0\}$, x_{0j} can be equal to 0,1 or 2. The latter case corresponds to a immediate return trip between the depot and customer j. Under these assumptions, the VRP formulation is as follows.

$$\min_{x} \sum_{[i,j]\in E} c_{ij}x_{ij} \tag{1.1}$$

subject to

$$\sum_{j\in V\setminus\{0\}} x_{0j} = 2m \tag{1.2}$$

$$\sum_{i<k} x_{ik} + \sum_{j>k} x_{kj} = 2 \qquad k \in V \setminus \{0\} \tag{1.3}$$

$$\sum_{\substack{i\in S, j\notin S \\ \text{or } i\notin S, j\in S}} x_{ij} \geq 2b(S) \qquad S \subset V \setminus \{0\} \tag{1.4}$$

$$x_{ij} = 0 \text{ or } 1 \qquad i,j \in V \setminus \{0\} \tag{1.5}$$

$$x_{0j} = 0 \text{ or } 1 \text{ or } 2 \qquad j \in V \setminus \{0\} \tag{1.6}$$

In the above formulation, the objective (**??**) is to minimize the total costs. Constraint (**??**) sets the degree of vertex 0 in accordance with the given number of vehicles m. It should be noted that if not known m can be a decision variable. Constraints (**??**) guarantee that two edges are incident to each customer vertex. The term $b(S)$ in constraints (**??**) is a lower bound on the number of vehicles required to visit all customers in S. The constraints force any subset of customers to be connected to the depot. Furthermore, constraints (**??**) ensure that the capacity restriction is preserved. One common way is to define $b(S)$ is $\left\lceil \frac{\sum_{i\in S} q_i}{Q} \right\rceil$. Constraints (**??**) and (**??**) guarantee integrality.

1.2. Time and Timing

The research presented in this thesis approaches the VRP both from a time and a timing perspective. Figure **??** describes the positioning of the research, with respect to these perspectives. The time perspective relates to the actual travel time of a vehicle, and thus relates to an operational cost. The timing perspective, relates to the customer service aspect of the problem. It mainly deals with arriving within a time window at a customer location.

Section **??** highlights the main time perspectives treated in Chapters 2 and 3. Section **??** presents the customers service concept used in Chapters 4 and 5.

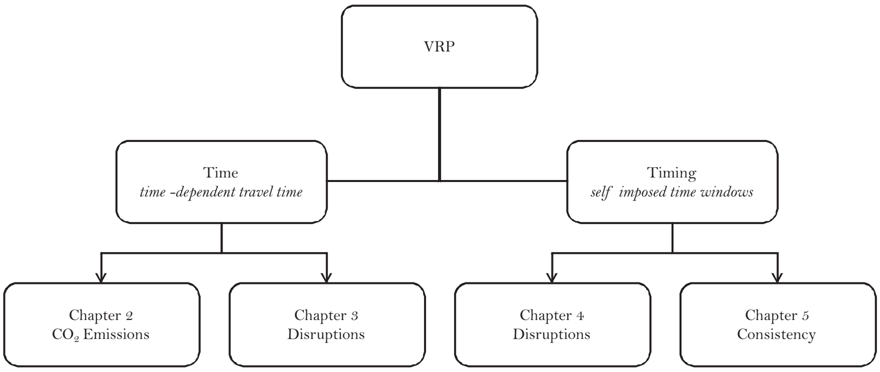

Figure 1.1 Overview of the Thesis

1.2.1 The time perspective

In the core of the Vehicle Routing Problem lies the objective of minimizing the total operational cost. The overwhelming majority of research considered this operational cost in terms of total travel time or total distance. In the VRP formulation (as presented in Section **??**), the total cost of a solution is simply the sum of the arc costs that comprise it. These costs are usually named distance or travel time interchangeably. The underlining equivalence between travel time and distance is due to the assumption that speed is constant throughout the planning period. Thus, the travel time is merely a linear transformation of the distance.

In reality, speeds are not constant throughout the day. There is a sizeable variability with respect to the time one travels when considering several starting times. **?** classified potential causes of variability in travel times into two main components. The first is attributed to random events such as accidents. The second is due to temporal variations resulting from hourly, daily, weekly or seasonal cycles in the average traffic volumes, i. e., traffic congestion. In recent years traffic congestion has been rising substantially. For example, **?** showed that in 2007 congestion in the USA caused 4.2 billion hours of delay. **?** demonstrated that recurrent occurrences during peak hours is the major part of traffic delay. In order to account for variations in travel times, a number of researchers studied the Time-Dependent Vehicle Routing Problem (TDVRP). In essence, the TDVRP relaxes the assumption of constant speed throughout the day.

In Chapters 2 and 3, time-dependent travel times are incorporated in the VRP. In these chapters time-dependent travel times are modeled similar to **?** and **?**. In the TDVRP, the operational cost is expressed exclusively in terms of total travel time, where the travel time between nodes is a function of both distance and starting time, i. e., given a distance its associated travel time is a function of the moment the vehicle starts traversing it. This implies that the decision variables need to account for the actual time an arc is to be traversed, along with the routing decisions.

1.2.2 The timing perspective

The Vehicle Routing Problem with Time Windows (VRPTW) extends the VRP by requiring that the service at each customer starts within a given customer time window. The different time windows for each customer are given as constraints to the problem. While arriving at the customer after the upper limit of its time window is prohibited, vehicles are allowed to wait at no cost if they arrive early. Some research considers soft time windows by which time window violations are permitted at a penalty cost. Time windows are seen as a customer service aspect in the VRP.

The VRPTW has been the subject of a great amount of research. **?** present an overview of exact methods for the VRPTW. Due to the VRPTW's high complexity coupled with its applicability to real life, much research focused on heuristic solution methods. **?** survey route construction and local search algorithms, while in **?** a survey of metaheuristics is presented. In most of the surveyed papers, the objective of the VRPTW was first to minimize the number of vehicles needed to serve all customers, and second to minimize the total travel times. Thus, the objective of the VRPTW remains minimizing operational cost subject to satisfying customer time window requirements.

Chapters 4 and 5 take a different standpoint with respect to customer service. Realizing that many carrier companies quote their expected arrival times to their customers, the concept of self-imposed time windows (SITW) is introduced. The notion of SITW reflects the fact that time windows are determined by the carrier company and not by the customers. However, once a time window is quoted to the customer the carrier company strives to provide service within this time window. SITW are fundamentally different from the time windows in the well-studied VRPTW. In the latter, the customer time windows are exogenous constraints, while SITW treat time windows as endogenous decision variables. Clearly, SITW give additional flexibility to the carrier company. This flexibility is likely to reduce operational costs when compared to VRPTW.

Considering SITW in VRP entails that along with the routing decisions timing decisions are to be taken as well. Given the above concepts the VRP with self-imposed time windows (VRP-SITW) is viewed as a problem in between the VRP and the VRPTW. The VRP accounts for no customer service aspects and optimizes exclusively on operational cost. The VRPTW is constraint by obeying given customer time windows. The VRP-SITW considers customer service with flexible time windows. Chapters 4 and 5 consider different realistic stochastic environments for embedding SITW.

1.3. Uncertainty in VRP

The basic formulation of the VRP assumes that all problem parameters are deterministic. However, in reality carrier companies are faced with various types of uncertainty. To encompass these uncertainties a number of stochastic elements have been introduced to the VRP. In many cases, stochastic VRP are cast within the framework of *a priori* optimization problems. An *a priori* solution attempts to obtain the best solution, over all possible problem scenarios, before the realization of any scenario (**?**).

The motivation behind using the *a priori* approach is that it is impractical to consider an a posteriori approach, that recomputes the optimal solution upon the realization of random event. The research presented in this thesis treats problems for which decisions are taken in an *a priori* manner. Chapters 3-5 consider stochastic events. These events are revealed only upon the execution of the routes. However, no *online* amendments to the solutions are made to accommodate these random events. The advantages of an *a priori* approach lay in the construction of robust solutions that are able to efficiently cope with random events. For an overview of stochastic vehicle routing the reader is referred to **?**. In this thesis, uncertainty in VRP is modeled by discrete distributions. We construct solutions that account for the expected costs of these events. We treat three types of uncertainties.

1. In Chapter 3, stochastic service times are considered. We model situations where the vehicle arrives at a customer location and the customer is not ready to receive the vehicle. In this situation, the vehicle waits until the customer is ready to receive service.

2. In Chapter 4, stochasticity in travel time is modeled to describe variability caused by random events such as car accidents or vehicle break down.

3. In Chapter 5, the objective is to construct a long term plan for providing consistent service to reoccurring customers. Thus, stochastic customers are considered.

1.4. Overview of the thesis

The thesis builds upon the established literature of the VRP and extends it to portray more realistic aspects. The presented research identifies relevant problems, establishes models that depict them and develops efficient solution frameworks that fit them.

Chapter 2 addresses, the ever growing, environmental concerns by incorporating CO_2 emissions related costs in the routing problem. We propose a framework for modeling CO_2 emissions in a time-dependent VRP context (E-TDVRP). As the generated amount of CO_2 emissions is correlated with vehicle speed, our model considers limiting vehicle speed as part of the optimization. The chapter explores the impact of limiting vehicle speed on CO_2 emissions. Furthermore, the trade-off between minimizing travel time or CO_2 emissions is analyzed.

The emissions per kilometer (as a function of speed) is a convex function with a unique minimum speed v^*. However, the results presented in Chapter 2 show that limiting vehicle speed to this v^* might be sub-optimal, in terms of total emissions. During congestion vehicles are constrained by lower speeds and produce much higher emissions. Consequently, in order to minimize emissions, it might be worthwhile to travel at speeds higher than v^* in order to avoid congestion. Furthermore, upper and lower bounds on the total amount of emissions that may be saved are presented.

Chapter 3 focuses on a problem expressed by a number of carrier companies. During the execution phase of vehicle routing schedules many unexpected delays are observed. These delays are attributed to the situation where customers are not ready to receive their goods. This problem is studied in conjuncture with time-dependent travel times. The construction of *a priori* solutions that minimize the damages inflicted by these disruptions is the objective of this chapter.

Chapter 3 also defines defines the Perturbed Time-Dependent VRP (P-TDVRP) model which is designed to handle unexpected delays at the nodes. The cost-benefit analysis of the P-TDVRP solutions emphasizes the added value of such solutions. Furthermore, the chapter identifies situations capable of absorbing delays. i. e., where inserting a delay leads to an increase in travel time that is less than or equal to the expected delay length itself. Based on this, we analyze the structure of the solutions resulting from the P-TDVRP.

Chapter 4 is centered around the problem of quoting service times to the customers. The problem arises when customers place orders beforehand and the carrier company has the control of determining a service time window. This time window is later communicated to the customers as the expected arrival time window of the carrier. Once a time window is quoted to a customer, the carrier company strives to respect it as well as possible. The time windows are treated as soft time windows, where penalties are incurred for late arrivals. The problem is optimized under the assumption that stochastic disruption may occur during travel times. Given a discrete distribution of disruptions, the objective is to construct routes that minimize the sum of operational and service costs. The latter corresponding to the penalties caused by late arrivals with respect to the quoted time window.

In Chapter 4, the notion of self-imposed time windows is defined and embedded in the VRP-SITW model. Two main types of experiments were conducted; one compares the VRP with VRP-SITW while the other compares VRPTW with VRP-SITW. The first set of experiments assesses the increase in operational costs caused by incorporating SITW in the VRP. The second set of experiments enables evaluating the savings in operational costs by using flexible time windows, when compared to the VRPTW.

Chapter 5 extends the timing dimension in the context of the consistent vehicle routing problem. **?** defined consistency as having the same driver visiting the same customers at roughly the same time on each day that these customers require service. This definition stems from the needs expressed by United Parcel Service (UPS). Consistent service facilitates building bonds between customers and drivers. Such bonds could be translated into additional revenues. Little work has been conducted on this problem. Moreover, the existing literature considers settings were full periodic knowledge is known about the problem. Chapter 5 broadens the definition of the problem to include fully stochastic parameters.

Chapter 5 models the consistent vehicle routing with stochastic customers (SConVRP). Driver consistency is forced by assigning a unique driver to visit a customer. The requirement that the customer is visited at the time same when he places an order is characterized as temporal consistency. To model temporal consistency the concept of SITW was used. Chapter 5 formulates the SConVRP and describes an exact solution approach.

1.5. Solution methods

In this section, we survey the relevant solution methods and we summarize the solution approaches adopted in each chapter.

The VRP is NP-hard since it contains the Traveling Salesman Problem (TSP) as a special case (where $m = 1$ and $Q = \infty$). In comparison, the VRP is considered to be harder than the TSP. Exact algorithms have been capable of solving TSP instances with thousands of vertices (**?**). Conversely, exact algorithms designed for the VRP can solve up to 100 customers (see, e. g., **?** for a survey). To encompass the complexities faced by carrier companies, a number of modifications to the original problem are proposed in the literature. These modifications entail additional constraints to the VRP, which in return generally increase the difficulty of the problems. Thus, the aspiration of solving more practical settings of the VRP mostly prompted the development of appropriate heuristics or exact methods, where possible.

Classical Heuristics are classified into three categories by **?**. *Constructive heuristics* that gradually build a feasible solution. Nodes are chosen based on some cost minimization criteria. One example is the savings heuristics proposed by **?**. The second category of heuristics is *two-phase heuristics*, where in the first phase customers are clustered into feasible routes and in the second phase the actual routes are established, e. g., the sweep algorithm **?**. The third category of heuristics is *improvement methods*. These are based on the concept of iteratively improving the solution to a problem by exploring neighboring ones (see, e. g., **?**). As mentioned by **?**, these classical heuristics have a large gap when compared to the best known solutions.

Over the years a number of *metaheuristics* have been applied to the VRP. **?** mention six main type of metaheuristics: simulated annealing, deterministic annealing, tabu search, genetic algorithms, ant systems, and neural networks. **?** mention that while the success of any method is related to its implementation features, it is fair to say that tabu search clearly outperforms competing approaches.

Tabu Search was first introduced by **?**. The procedure explores a solution space at each iteration as it moves from one solution to another. Given a solution x_t at iteration t a subset neighborhood $N(x_t)$ of x_t is explored and the best solution found in this neighborhood is set as x_{t+1}. By this definition, the procedure allows non-improving moves. Thus, to avoid, cycling solutions that were previously explored are forbidden, tabu, for a number of iterations. Over the years Tabu Search has been applied successfully to the VRP, we mention the work of **?**, **?**, and **?**.

We adopted Tabu Search in Chapters 2-4. In Chapter 2, we integrated time-dependent travel time in the search. Furthermore, the vehicle speed limit, which is a decision parameter is embedded in the solution procedure. In Chapter 3 perturbations are imposed, depicting disruptions in service times. To our knowledge, the proposed solution procedure, in Chapter 3, is the first application of the methodology proposed by **?** to a discrete problem.

In Chapter 4, we propose a heuristic solution approach combining an LP model in a Tabu Search framework. The tabu search assigns customers to vehicles and establishes the order of visit of the customers per vehicle. Detailed timing decisions are subsequently generated by the LP model, whose output also guides the local search in a feedback loop.

Chapter	CO_2 Emissions	Time-Dependent Travel Time	Disruptions	Self-Imposed Time Windows	Stochastic Customers
2	X	X			
3		X	X		
4			X	X	
5				X	X

Table 1.1 Thesis outline: main features by chapter

In Chapter 5, a stochastic programming formulation is presented for the consistent VRP with stochastic customers. This differs from the models proposed in the literature which assume known customer demands for a given period. An exact solution method is proposed by adapting the 0-1 integer L-shaped algorithm, introduced by **?**. The proposed algorithm is based on the L-shaped method proposed by **?** for continues problems and on Benders decomposition (**?**). The 0-1 integer L-shaped algorithm solves problems with binary first stage variables and integer recourse. It applies branch-and-cut to a relaxed version of the problem. The algorithm introduces lower bounding functionals and optimality cuts to the current problem and converges with a finite number of steps.

The 0-1 integer L-shaped algorithm has been applied to a number of stochastic VRP variants. **?** adapted the algorithm for the single VRP with stochastic demand. **?** solved the VRP with stochastic demand. Finally, **?** applied the algorithm to the VRP with stochastic customers and demand.

1.6. Outline of the Thesis

The general features of each of the chapters are summarized in Table **??**. The time dimension is a key component in Chapters 2 and 3. Including customer service is a focal element in Chapters 4 and 5. Chapter 2 considers the VRP with CO_2 emissions along with TDVRP. Chapter 3 incorporates disruptions in the TDVRP. Chapter 4 optimizes the VRP with SITW under disruptions. Finally, Chapter 5 solves the consistent VRP with stochastic customers and makes use of SITW to guarantee temporal consistency. The main conclusions are discussed in Chapter 6. The relevant literature is surveyed in each chapter separately.

Chapter 2

Analysis of Travel Times and CO_2 Emissions in TDVRP

The ever growing concern over Greenhouse Gases (GHG), has led many countries to take policy actions aiming at emissions reductions. Most notable is the Kyoto Protocol which enforces countries to reduce a basket of the six major GHG by the year 2012 by 5.2% on average (compared to their 1990 emission levels). Next to this, a number of other initiatives have emerged to particularly control CO_2 emissions. For example, more than 12,000 industrial plants in the EU are subjected to CO_2 caps, enabling the trade of emissions rights between parties. For a comprehensive survey of GHG trade market models, the reader is referred to **?**.

The importance of environmental issues is continuously translated into regulations, which potentially have a tangible impact on supply chain management. As a consequence, there has been an increasing amount of research on the intersection between logistics and environmental factors. **?** identified potential combinatorial optimization problems where Green Logistics is relevant. **??** discussed the integration of environmental management in operations management. **?** did so in the context of sustainable operations management.

European road freight transport uses considerable amounts of energy (**?**). **?** note that predictions are that the UK will meet the Kyoto targets. However, they highlight that within the period from 1985 until 1995, energy use across all sectors grew only by 7% while transport energy use grew by 31%. Similar findings were observed by **?**. They state that in the period 1991-2001 road freight traffic in Germany increased by 40%. Moreover, in 2001 traffic was responsible for about 6% of total CO_2 emissions in Germany. More substantial CO_2 increases are observed in India by **?**, where the CO_2 emissions from road transport in the year 2000 have increased by almost 400% compared to 1985. In the context of examining scenarios for GHG, **?** mentioned that in California the transportation sector accounted for over 40% of GHG for 2006, making it by far the largest contributor. **?** acknowledge that new vehicle designs are more efficient in terms of emissions. However, this is outweighed by the influx in transport growth rate in the EU. **?** mention that CO_2, which is directly related to the consumption of carbon-based fuel, is regarded as one of the most serious threats to the environment through the greenhouse effect. Globally, transportation accounts for about 21% of CO_2 emissions. One of the main drives of this trend is the increasing demand for goods transport. For example, in Europe the total road freight transport volume increased by 43% between 1992 and 2005 (**?**). CO_2 is identified as the most important greenhouse gas in the Netherlands, as it accounted for 80% of total emissions in 1995 (see **?**). Thus, there is a clear necessity to control the CO_2 emissions produced.

Road transport accounts for a large portion of the CO_2 emissions, of which goods transport constitutes a sizeable portion. Thus, there is a need for addressing environmental concerns. Carrier companies may voluntarily adopt green policies if this is aligned with profitability. This could be in the form of GHG trading mechanisms, or when CO_2 becomes a taxable commodity. Another reason for adopting green policies is the marketing potential of a greener company image, e.g., controlling the carbon footprint. Furthermore, new regulations might force companies to change practices. It is worth mentioning that the department for environmental food and rural affairs in the UK values the social cost of carbon between £35 and £140 per tonne. In essence, pricing carbon emissions leads to an assessment of its economic impact and regulations might be formed accordingly. In conclusion, either for exogenous or endogenous reasons, change is anticipated in transportation management with respect to environmental factors. We argue that Logistics Service Providers should contemplate on how to deal with these issues.

The focus of this chapter is on incorporating CO_2-related considerations in road freight distribution, specifically in the framework of Vehicle Routing Problems (VRP). Both CO_2 emissions and fuel consumption depend on the vehicle speed, which changes throughout the day due to congestion. Thus, it is very relevant to study the problem on-hand in conjunction with time-dependent travel times, i.e., where the travel time depends on the time of day at which a distance is traversed. Time dependency is modeled by partitioning the planning horizon into free flow speed periods and periods with congestion, i.e., lower speeds. We introduce a new variant of the VRP that accounts for travel time, fuel, and CO_2 emissions costs. This results in the Emissions based Time-Dependent Vehicle Routing Problem, denoted by E-TDVRP.

The E-TDVRP builds on the possibility of carrier companies to limit the speed of their vehicles. Thus, the vehicle speed limit is explicitly part of the optimization. The traditional time-dependent vehicle routing problem (TDVRP) optimizes exclusively on travel times and thus does not consider limiting vehicle speed. However, we show that when accounting for fuel, travel time, and emissions, controlling vehicle speed is desirable from a total cost point of view.

The E-TDVRP is clearly a complex problem, as in addition to the complexity of the (TDVRP), it also determines the free flow speed. This implies that one needs to allocate customers to vehicles, determine the exact order which customers are visited and set the free flow speed permitted. The free flow speed impacts the resulting travel time functions of each arc, and in return, affects the moment vehicles go into congestion. We assume that the congestion speed remains constant, as it is imposed by traffic conditions. This leads to a situation where speeds are controlled in particular time periods of the day, i.e., exclusively for free flow speed. The E-TDVRP can be reduced to two subproblems: one where only the CO_2 emissions are taken into account in the optimization (i.e., a pure environmental model) and one where only travel times are considered (i.e., a pure logistical cost model). As such, we study the trade-off between minimizing CO_2 emissions as opposed to minimizing total travel times. In addition, we develop bounds for the potential reduction in CO_2 emissions. These bounds are based on solutions of the standard time-independent VRP. These bounds aid decision makers in evaluating the maximum reduction in emissions, since most industrial optimization tools consider time-independent travel times.

The remainder of the chapter is organized as follows, Section **??** reviews the relevant literature. Section **??** describes the E-TDVRP model. It also introduces bounds for the potential savings in CO_2 emissions. The solution methods and the experimental settings are discussed in Section **??**. The results are presented in Section **??**. Finally, Section **??**, highlights the main findings and indicates directions for future research.

2.1. Literature Review

? highlighted the value of traffic flow information related to emissions. Their results showed that calculating emissions under constant speed assumptions can be misleading, with differences of up to 20% in CO_2 emissions on an average day for gasoline vehicles and 11% for diesel vehicles. During congested periods of the day these difference rose up to 40%. Similar results were shown by **?**, on a number of roads. Such results are motivated mainly by the fact that CO_2 emissions are proportional to fuel consumption (**?**), and thus are speed-dependent. The relation between emission values and vehicle speed leads to the study of the VRP problem in a time-dependent framework. In what follows, we first discuss the literature concerning the TDVRP, and then review the literature dealing with routing and emissions.

Assigning and scheduling vehicles with a limited capacity, to service a set of clients with known demand, is a problem faced by numerous carrier companies. It has been extensively researched in the literature as the well-known vehicle routing problem (VRP). For a comprehensive review, the reader is referred to **?**. In its most standard version, the problem deals with minimizing costs subjected to satisfying customer demand, while vehicle capacity constraints are maintained. Each customer is visited by a single vehicle, each vehicle starts and ends its route at the depot. The relevance of VRP to real-life practice coupled with the hard nature of the problem has attracted much research. To model reality more accurately, numerous features have been incorporated in the problem. One of which is including speed changes over the day, in an attempt to account for traffic congestion experienced in certain periods of the day.

Incorporating time-dependent travel times between customers in the VRP has been adopted recently by a number of researchers. The objective of the TDVRP, in most cases, is similar to that of the VRP, i.e., minimizing costs (**?**). However, in the TDVRP the travel time cost depends on the time of day a distance is traversed. Modeling time dependency is mostly done by associating different speeds to a number of time zones within the planning horizon. **?** motivate variability in travel time by random events, such as accidents and cyclic temporal variations in traffic flow. They proposed a mixed integer programming approach to the TDVRP. **?** proposed a dynamic programming formulation to solve the time-dependent TSP. Both these papers modeled travel times by discrete step functions where travel times are associated with different time zones. While this modeling approach does capture variability in travel times, it enables the undesired effect of surpassing. This effect implies that a vehicle departing at a certain time might surpass another vehicle that started traveling earlier. This limitation was discussed by **?**, **?**, **?**, and **?**. All these papers model time dependencies complying with the FIFO (First-In First-Out) assumption, which does not allow for surpassing. This is done by using appropriate piecewise linear functions for travel times. In this chapter, we adopt travel time functions similar to the ones used by **?**, and thus adhere to the FIFO assumption.

? modeled the TDVRP by partitioning the day into three speed zones, where the speed differences due to congestion were determined by different factors of the free flow speeds. The travel time profiles were constructed by piecewise linear functions. **?** discuss a general framework for integrating time dependent travel times in a number of VRP routing algorithms. Furthermore, they provide an overview of traffic information systems from which data can be collected. Modeling time-dependent travel times would benefit from these data. Based on empirical traffic data, queueing models were developed by ? to model congestion, where the model parameters were set to incorporate different traffic flows and weather conditions. The analysis of the data resulted in average speeds for different time zones (see also ?). These speeds where later used in ? to model the TDVRP, which was solved by means of Tabu Search. ? proposed a modified genetic algorithm to solve the TDVRP.

Not much research has been conducted on the VRP under minimizing emissions. ? studied the environmental impact of grocery home delivery. This however was done by converting distance into emissions, irrespectively of speed changes. The effect of speed changes is incorporated in ?. He studied emissions in the context of grocery home delivery vehicles, where real traffic data was used to derive fuel consumption and emissions. A similar detailed methodology was used by ?. Both ? and ? considered CO_2 minimization as an optimization criteri. In a more aggregate view, ? presented an emission-based trip assignment optimization model. They also explored potential emission reduction under the assumption that drivers choose emission minimizing routes. Assessing vehicle emissions can be very complex, as emissions depend on factors such as the age of the vehicle, engine state, engine size, speed, type of fuel and weight (?). For our study, we used speed emission functions from the MEET model (European Commission ?). In this chapter, we focus on CO_2 and leave other pollutants for further research.

2.2. Description of E-TDVRP

This section starts with a complete description of the E-TDVRP model and details the most relevant parts of this model. Section ?? elaborates on the computation of the travel times. Section ?? explains the computation of the CO_2 emissions and fuel consumptions.

In general, the vehicle routing problem (VRP) is described by a complete directed graph $G = (V, A)$ where $V = \{0, 1, \dots, n\}$ is a set of vertices and $A = \{(i, j) : i <> j \in V\}$ is the set of directed arcs. The vertex 0 denotes the depot; the rest of the vertices represent customers. For each customer, a non-negative demand q_i is given ($q_0 = 0$). A non-negative cost c_{ij} is associated with each arc (i, j). The objective is to find, for a given number of vehicles N, the minimum costs where the following conditions hold: every customer is visited exactly once by one vehicle, all vehicle routes start and end at the depot, and every route has a total demand not exceeding the vehicle capacity Q. This definition is valid for the E-TDVRP model presented in this chapter.

The E-TDVRP differs from the VRP as it considers time-dependent travel times.

Furthermore, it considers fuel and emission costs. The objective function of the E-TDVRP is the sum of all these costs. Limiting the free flow speed of the vehicles is examined in this problem. Thus, in addition to the routing solution the E-TDVRP has a decision variable v_f corresponding to the upper limit of the vehicle speed. The speed limit v_f is imposed on all vehicles, i.e., all routes have the same limit. Thus, v_f is considered as a tactical choice of the carrier company.

We define a solution as a set S with s routes $\{R_1, R_2, ..., R_s\}$ where $s \leq N$, $R_r = (0, .., i, ..., 0)$, i.e., each route begins and ends at the depot. We write $i \in R_r$, if the vertex $i \geq 1$ is part of the route R_r (each vertex belongs to exactly one route). We write $(i, j) \in R_r$, if i and j are two consecutive vertices in R_r. An E-TDVRP solution is defined by the solution set S coupled with a vehicle speed limit v_f.

A solution (S, v_f) results in travel times $TT(S, v_f)$ and emissions $E(S, v_f)$. The computation of $TT(S, v_f)$ is explained in section **??**. The computation of $E(S, v_f)$ is discussed in section **??**.

We define three cost factors in the E-TDVRP: the hourly cost of driver (with a cost of α €/hr), the cost of fuel (costing β €/liter) and the cost of CO_2 emissions (γ €/kg). Since fuel consumption is directly related to CO_2 emission, we use the factor of h (equal to $\frac{1}{2.7}$ liter/kg) for converting CO_2 emissions into fuel consumption, similar to **?**. The objective function for E-TDVRP is given by Equation (**??**).

$$F(S, v_f; \alpha, \beta, \gamma) = \alpha TT(S, v_f) + (\beta h + \gamma) E(S, v_f) \tag{2.1}$$

The first part of the objective function considers the travel time costs. The second part considers the composite costs combining fuel and emissions. We explicitly consider both, as the cost parameters for fuel (β) and emissions (γ) are different.

The E-TDVRP can be reduced into two special cases. Setting α to zero the model minimizes solely on the costs of fuel and CO_2, which is equivalent to minimizing $E(S, v_f)$, with an objective function value $F(S, v_f; 0, \beta, \gamma)$. Similarly, setting both β and γ to zero results in a model that minimizes $TT(S, v_f)$, with an objective function value $F(S, v_f; \alpha, 0, 0)$. These special cases facilitate a trade-off analysis showing the additional travel time required to obtain minimal CO_2 emissions.

For completeness, we summarize the speed-related notation we will use in the reminder of this section:

- v_f: The speed limit set in E-TDVRP.
- v_c: The congestion speed imposed by traffic.

- v^*: The speed that yields the optimal CO_2 emission per km value.
- v_f^*: The optimal speed for $E(S, v_f)$.
- v_h: The upper bound value of v_f^*.
- v': The speed used for computing the upper bound on $E(S, v_f)$.

2.2.1 Determining $TT(S, v_f)$

The time-dependency of travel times throughout the planning horizon, in essence, is driven by changing speeds in different time zones. We subject all links to given speed profiles. This stems from the notion that most motor-ways, on average, follow the same pattern of having a morning and evening congestion periods (see also **?** for a similar reasoning). Moreover, collecting data for each link and each time zone is infeasible from an operational standpoint.

Figure **??** provides an illustration of how speeds are translated into travel time profiles. The left side depicts a speed profile that starts with a congestion speed v_c until time a. After time a the vehicle can travel at free flow speed v_f. Subjecting a given distance d to the speed profile on the left side of the figure, results in the travel time profile on the right side of Figure **??**. Note that while the x-axis for the speed profile is the time of day, the x-axis for the travel time figure is starting times. For a given distance d and for every starting time the figure on the right provides the travel time. The main intuition for modeling the profile is that during the first period (up to $a - TT_c$) it takes TT_c time units to traverse d. Since, throughout that period the vehicle will be driving with speed v_c along the entire link. However, starting from time $a - TT_c$ up to point a, the vehicle will be in the transient zone, where the vehicle will start traversing part of the link with speed v_c and the remainder with a speed v_f. Starting at point a the speed will remain constant at v_f with the travel time equaling TT_f. Thus, the travel time is a continuous piecewise linear function over the starting times.

The linearity in the transient zone stems from the stepwise speed change which imposes different speeds over time. The slope can be defined as $\frac{TT_f - TT_c}{TT_c}$. By substituting TT_c with $\frac{d}{v_c}$ and TT_f with $\frac{d}{v_f}$, we obtain that the slope is equal to $\frac{v_c - v_f}{v_f}$, which is independent of d. However, the intersection with the Travel Time axis is a function of the distance and it is equal to $\frac{(v_f - v_c)}{v_f}a + \frac{d}{v_f}$. Therefore, the travel time function between customers i and j depends on the distance d_{ij} between these customers. We define $g(t)$ as the travel time function associated with any starting time t for Figure **??**.

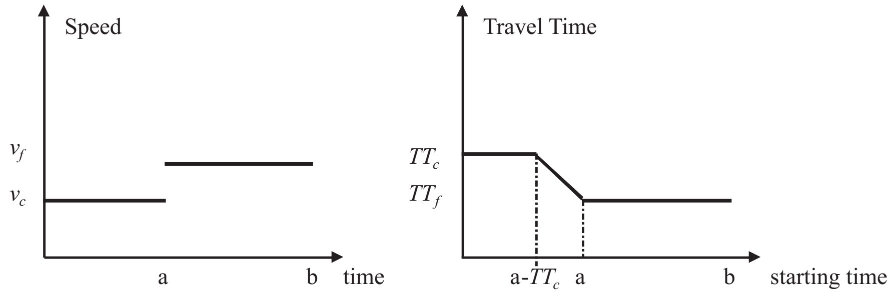

Figure 2.1 The conversion of speed into travel times

$$g(t) = \begin{cases} TT_c & \text{if } t \leq a - TT_c \\ \frac{v_c - v_f}{v_f} t + \frac{(v_f - v_c)}{v_f} a + TT_f & \text{if } a - TT_c < t < a \\ TT_f & \text{if } t \geq a \end{cases}$$

For starting times within $(0, a - TT_c)$ FIFO is satisfied since the speed associated with these starting times is constant, given two starting times within this interval t and $t+\Delta$, both will arrive after TT_c time units from their departure. Similarly, FIFO holds for starting times after a. The line in the transient zone $(a - TT_c, a)$ can be viewed as a consequence of the FIFO assumption as well. Given two starting times t and $t+\Delta$, both in the transient zone, the arrival times are $t+g(t)$ and $t+\Delta+g(t+\Delta)$ respectively, the difference between the arrival times of $t+\Delta$ and t is $\Delta + \frac{v_c - v_f}{v_f}\Delta > 0$. And thus, the FIFO assumption holds in the transient zone. Note that in a speed-decreasing situation the FIFO assumption holds even for step travel time changes. Nonetheless, for consistency, speed drops are constructed exactly in the same manner as speed increases. The construction of travel times is equivalent to integrating the distance over the different speeds. The proposed model is convenient since it requires speed values and the point in time in which the speed changes. Let $t(r) \in \{t_1, ..., t_k\}$ denote the starting time of route R_r. Let $t(r, i)$ denote the departure time at node i, for route R_r. $TT_{i,j,t(r,i)}$ represents the travel time between node i and node j when starting at time $t(r, i)$. Equation (**??**) considers the total travel time associated with a solution S and v_f.

$$TT(S, v_f) = \sum_{r} \sum_{t(r)} \sum_{(i,j) \in R_r} TT_{i,j,t(r,i)} \tag{2.2}$$

2.2.2 Determining $E(S, v_f)$

In this chapter, we use emission functions provided in the MEET report (European Commission **?**). The function $\theta(v)$ provides the emissions in *grams* per kilometer for speed v. Equation (**??**) depicts the amount of emissions per km, given that a vehicle is at speed v.

$$\theta(v) = K + av + bv^2 + cv^3 + d\frac{1}{v} + e\frac{1}{v^2} + f\frac{1}{v^3} \tag{2.3}$$

The coefficients $(K, a, \ldots, f)$ differ per vehicle type and size. Here, we focus on heavy duty trucks weighing 32-42 tons. The coefficients for the CO_2 emissions for this specific vehicle category are $(K, a, b, c, d, e, f) = (1576, -17.6, 0, 0.00117, 0, 36067, 0)$. Figure **??** depicts this CO_2 (kg/km) emissions function. This function has a unique minimum, we define v^* as the integer speed which achieves this minimum ($v^* = 71$ km/hr).

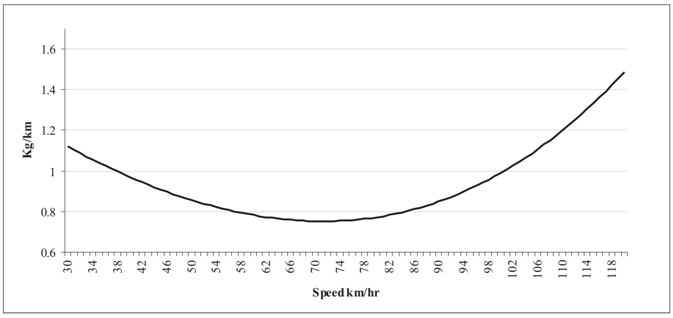

Figure 2.2 CO_2 emissions versus speed

We denote the amount (in grams) of CO_2 emissions produced by traversing arc (i, j) at time t with a free flow speed v_f by $E_{i,j,t(r,i)}(v_f)$. The average speeds are used to calculate the emissions per km by Equation (**??**). Multiplication of emissions per km times distances traversed yields the total amount of CO_2 emissions. We define $\Omega[d_{ij}, t(r, i), v_f]$ as the average speed for traversing arc (i, j) when leaving i at time $t(r, i)$, with speed limit v_f. Consequently, $E_{i,j,t(r,i)}(v_f)$ is given by Equation (**??**).

$$E_{i,j,t(r,i)}(v_f) = \theta\left(\Omega[d_{ij}, t(r, i), v_f]\right) d_{ij} \tag{2.4}$$

The objective function for term $E(S, v_f)$ is defined in Equation (**??**) summing the total CO_2 emissions produced by the different routes R_r in solution S with speed limit v_f.

$$E(S, v_f) = \sum_r \sum_{(i,j)\in R_r} E_{i,j,t(r,i)}(v_f) \tag{2.5}$$

Both from a practical and an optimization point, we choose to work with integer speeds. Optimizing on $E(S, v_f)$ implies then finding an optimal integer free flow speed applied to all routes R_r in S. The dependence of $E_{i,j,t(r,i)}$ on the speeds associated with d_{ij} makes the free flow speed v_f a decision variable. We emphasizes that limiting the speed of a vehicle means that in free flow the vehicle's speed is limited. However, in congestion the vehicle is restricted by the congestion speed. In essence, while congestion zones can be seen as constraints, in free flow zones the maximum speed is decided upon.

Modifying the speed profile affects the travel time profiles. As explained in Section **??**, the change of speed at point a starts affecting the travel times already at point $a - TT_c$. If there is a speed drop at point a from a certain free flow speed v_f to a congestion speed v_c, travel times for a given distance d will start to increase if it is traversed after time $a - \frac{d}{v_f}$. Altering the free flow speed will affect the travel time profile in a way that, if decreased, the vehicle will start experiencing congestion at earlier starting times. We demonstrate the effect of free flow speed on the total emissions produced with the example depicted in Figure **??**. Let $v_c = 40$ km/hr and consider two options for free flow speed: $v_1 = 71$ km/hr, and $v_2 = 72$ km/hr, for traversing a distance of 400 km with starting time 200 and $a = 500$ (in minutes). Setting the free flow speed to v_1 produces 2.8kg more CO_2 emissions than setting it to v_2. Moreover, setting the free flow speed to v_1 results in an increase of 15 minutes in travel times, with respect to v_2.

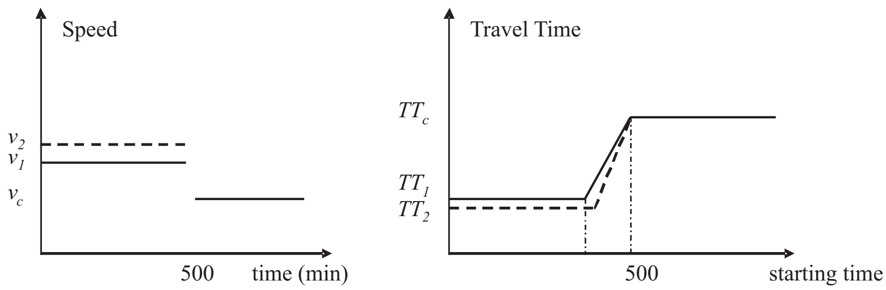

Figure 2.3 Example for two free flow speeds

Considering optimizing on $E(S, v_f)$, i.e., only on the amount of emissions, let the optimal speed be v_f^*. Proposition **??** shows that only a subset of potential speeds need to be taken into account for the $E(S, v_f)$ optimization. This proposition makes it possible to considerably reduce the search space for this problem.

Proposition 2.1 Given $v_c < v^*$, there exists a $v_h > v^*$ such that $\theta(v_c) = \theta(v_h)$ and $v^* \leq v_f^* \leq v_h$.

For the case of $\theta(v)$ convex and having a unique minimum, as in the case of the CO_2 emission function considered, Proposition **??** is straightforward to show. Since v_c is smaller than v^* there exists another speed v_h which is higher than v^* such that $\theta(v_c) = \theta(v_h)$. We show that $v_c < v_f^* < v_h$, for the setting in question.

Proof: Arguing by contradiction, assume that $v_f^* > v_h$. The shape of the emission per km function implies that $\theta(v_f^*) > \theta(v_h) = \theta(v_c)$. In such a case, being in free flow zones, i.e., where the speed is set to v_f^*, will produce higher emissions than in congestion. Any $\tilde{v} \in [v_c, v_h]$ will produce less emissions since $\theta(\tilde{v}) < \theta(v_f^*)$ and thus v_f^* is not optimal.

Again by contradiction, assume the $v_c < v_f^* < v^*$. For any $\widetilde{v} \in [v_c, v^*]$ there exists a $\hat{v} > v^*$ such that $\theta(\hat{v}) = \theta(\widetilde{v})$. Since for $\hat{v}$ the total time spent in congestion will be lower than for $\widetilde{v}$, we conclude that $v^* < v_f^* < v_h$. □

As previously mentioned, we only consider integer values for speeds. Thus, v_h is rounded up to the nearest integer value.

2.2.3 Bounds for $E(S, v_f)$

The $E(S, v_f)$ on its own is a difficult problem to handle: one needs to find a solution in a time-dependent setting in combination with setting the free flow speed. In this section we present bounds using the standard VRP (i.e., time-independent). Throughout the years various solution techniques have been developed to tackle this problem. Moreover, from a practical standpoint most carrier companies developed software solutions for the standard VRP. Hence, bounds based on these basic VRP solutions are easily implemented in practice too.

The **lower bound** is realized by traversing the total distance in the VRP solution with a speed v^*, since this implies the minimum distance traversed using the optimal emissions speed. In essence, the **upper bound** can be constructed by subjecting the VRP solutions to a speed Profile with a single speed drop, similar to Profile I in Figure **??**. However, our TDVRP setting is similar to Profile II in Figure **??**. For a given distance to be traversed starting at time zero, optimizing CO_2 emissions under Profile II would generate results superior or equivalent to the same optimization under profile I. For minimizing emissions, the optimal performance under profile I is an upper bound on the performance of Profile II. Thus, we construct an upper bound on the total CO_2 emissions by assuming a single speed decrease which is outperformed by a profile consisting of a speed drop followed by a speed increase.

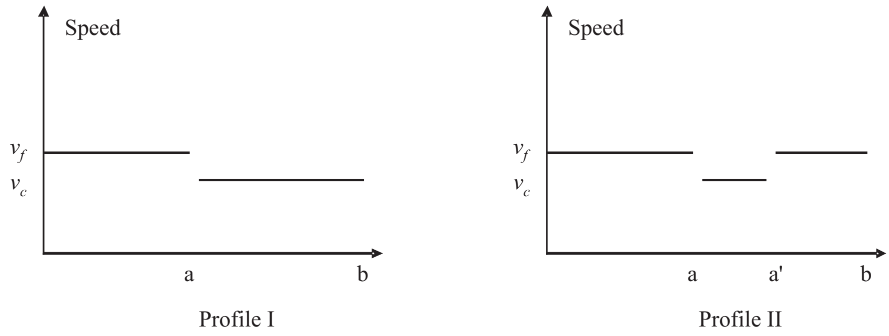

Figure 2.4 Speed profiles for the bounds on E-TDVRP

We analyze the case of subjecting the VRP solutions to Profile I. The VRP solution is a sequence of customers visited. We denote the distance of an arbitrary route by d. We distinguish two cases:

i) If the vehicle is not fast enough to avoid congestion, i.e., $v < d/a$, then the total emissions are given by the emissions in free flow $\theta(v)av$ plus the emissions in congestion which are $\theta(v_c)(d - av)$;

ii) If the vehicle travels fast enough to avoid congestion, i.e., $v \geq d/a$, then the total emissions would simply be given by $\theta(v)$ times the covered distance d;

According to the above distinction, define $E_1(d, v)$ and $E_2(d, v)$ as follows:

$$E_1(d, v) = \theta(v)av + \theta(v_c)(d - av)$$
$$E_2(d, v) = \theta(v)d.$$

Then, total emissions as a function of distance d, can be written as a function of the speed as:

$$E_d(v) := \begin{cases} E_1(d, v) & \text{if } v \in (0, \frac{d}{a}] \\ E_2(d, v) & \text{if } v \in [\frac{d}{a}, +\infty) \end{cases}$$

In our case $E_1(d, v)$, $E_2(d, v)$, and $\theta(v)$ are convex and have a unique minimum. The problem of finding an optimal speed involves finding the minimum of E_d:

$$\min \left\{ \min_{v \in (0, \frac{d}{a}]} E_1(d, v), \quad \min_{(v \in [\frac{d}{a}, +\infty)} E_2(d, v) \right\}. \tag{2.6}$$

Lemma 2.1 There exists a universal speed $v' > v^*$ such that for all $a, d > 0$ there exists a unique solution v_f^* to problem (**??**) given by:

$$v_f^* = \begin{cases} v^* & \text{if} \quad d \leq av^* \\ d/a & \text{if} \quad av^* < d < av' \\ v' & \text{if} \quad d \geq av'. \end{cases}$$

Proof: $E_1(d, v)$ has a unique optimal v'. This can be easily seen by derivation

$$\frac{\partial E_1(d, v)}{\partial v} = a(\theta'(v)v + \theta(v) - \theta(v_c)),$$

As earlier observed $E_2(d, v)$ has a unique optimal speed v^*. Furthermore $E_1(d, d/a) = E_2(d, d/a)$, so that E_d is continuous. Now, since by convexity we know that

- E_1 is decreasing for $v < v'$ and increasing for $v > v'$,
- E_2 is decreasing for $v < v^*$ and increasing for $v > v^*$,

it is sufficient to put together E_1 and E_2, distinguishing between the following three cases.

Case (i) : $d \leq av^*$.

$$E_d(v) = \begin{cases} E_1(v), & \text{decreasing} \quad \text{if } v \leq d/a \\ E_2(v), & \text{decreasing} \quad \text{if } d/a < v < v^* \\ E_2(v), & \text{increasing} \quad \text{if } v^* \leq v \end{cases}$$

so the minimum value is reached when $v = v^*$.

Case (ii) : $av^* < d < av'$.

$$E_d(v) = \begin{cases} E_1(v), & \text{decreasing} \quad \text{if } v \leq d/a \\ E_2(v), & \text{increasing} \quad \text{if } d/a < v \end{cases}$$

so the minimum value is reached when $v = d/a$.

Case (iii) : $av' \leq d$.

$$E_d(v) = \begin{cases} E_1(v), & \text{decreasing} \quad \text{if } v \leq v' \\ E_1(v), & \text{increasing} \quad \text{if } v' < v < d/a \\ E_2(v), & \text{increasing} \quad \text{if } d/a \leq v \end{cases}$$

so the minimum value is reached when $v = v'$. □

The importance of lemma **??** is that for large enough distances, $d \geq av'$ there exists a single speed v' which is optimal under Profile I. We use v' to construct the upper bound on $E(S, v_f)$.

Let S' be the optimal solution for the standard VRP. Let d'_i be the distance of route R_i in S'. We arrange the different routes in descending order, with respect to their distance, $S' = \{d'_{[1]}, ..., d'_{[s]}\}$. Equation (**??**) represents a lower bound on the amount of CO_2 emissions produced. The lower bound assumes that all distances can be traversed in v^*, so that absolute minimum emissions can be achieved.

$$\sum_{i=1}^{s} d'_{[i]}\theta(v^*) \leq E(S, v_f) \tag{2.7}$$

The upper bound is constructed by imposing Profile I with $v_f = v'$ onto S'. We define w as the largest index such that $d_{[w]} > av'$. Thus, routes $\{d'_{[1]}, ..d'_{[w]}\}$ will run into congestion. However, the routes $\{d'_{[w+1]}, .., d'_{[s]}\}$ will not suffer congestion. Equation (**??**) depicts this upper bound.

$$\begin{aligned} E(S, v_f) &\leq \theta(v')wav' + \theta(v_c)\sum_{i=1}^{w}(d'_i - av') + \theta(v')\sum_{i=w+1}^{s} d'_i \\ &\leq [\theta(v') - \theta(v_c)]wav' + \theta(v_c)\sum_{i=1}^{w} d'_i + \theta(v')\sum_{i=w+1}^{s} d'_i \end{aligned} \tag{2.8}$$

Given the standard VRP solution, these bounds, are rather straightforward to calculate, and enable decision makers to easily assess the maximal reduction in CO_2 emissions. Furthermore, these bounds are used to validate the proposed solution procedure.

2.3. Solution Method

We propose a Tabu search procedure for the E-TDVRP. Tabu search was first introduced by **??**. It makes use of adaptive memory to escape local optima. The method has been extensively used for solving the VRP, see **??** and **?** for examples. Tabu search is also used to deal with the time-dependent version of the VRP (see **?**,**?** and **?**). The E-TDVRP model differs from the TDVRP as it includes the free flow speed v_f as a decision variable. We chose to adapt a Tabu Search procedure to fit the E-TDVRP. Essentially, the procedure works on local search principles while updating the v_f. The algorithm searches a neighborhood with v_f. After a number

Algorithm 1 E-TDVRP algorithmic structure

```
construct initial solution S_0 with v_f = v_0 and compute F(S_0, v_f; α, β, γ)
si = σ
for i < I_max do
  generate the neighborhood of S
  choose the solution that minimizes F_2(S, v_f; α, β, γ) and is not tabu or satisfies
  its aspiration criteria
  update the tabu list
  if S is feasible and it is better than the current best feasible solution then
    update the best feasible solution
    set si = i + φσ
  end if
  if S is not feasible and is better than the current best solution then
    update the best solution
  end if
  update w
  if i > si then
    set current solution to best solution
    check for best speed within [v_f − τ, v_f + τ] and update v_f accordingly
    set si = i + σ
  end if
  if best solution found and I_max − i < ψ then
    set I_max = i + 2ψ
  end if
end for
return the best feasible solution
```

of iterations it checks the performance of the best solution so-far for different integer values of v_f. If a value which yields better results is found, the speed is updated and the search continues. Next, we describe the main components of the algorithm. The overall procedure is described in pseudo-code in Algorithm **??**.

The initial solution Z_0 is the output of a nearest neighbor heuristic with $v_f = v_0$. A neighborhood is evaluated by considering all possible 1-interchanges as proposed by **?**. The (0,1) interchanges of node $v_r \in R_r$ to route R_p are considered only if R_p contains one of η nearest neighbor of v_r. The (1,1) interchanges between nodes $v_r \in R_r$ and $v_p \in R_p$ are considered if R_p contains one of η nearest neighbor of v_r and R_r contains one of η nearest neighbor of v_p. After completing the neighborhood search, a 2-opt intra-route move is executed for each altered route. If node v_r is removed from R_r, reinserting v_r back into R_r is tabu for ℓ iterations. ℓ is randomly chosen between

$[\ell_{min}, \ell_{max}]$. However, we use an aspiration criterion similar to ?. This criterion revokes the tabu status of a move if it yields a solution with lower costs.

Similar to ?, we allow demand-infeasible solutions, i.e, routes with total demand exceeding the vehicle capacity. Such infeasible solutions are penalized, in proportion to the capacity violation, by the following objective function, extending $F(S, v_f; \alpha, \beta, \gamma)$:

$$F_2(S, v_f; \alpha, \beta, \gamma) = F(S, v_f; \alpha, \beta, \gamma) + w \sum_{R \in Z} \left[\left(\sum_{i \in R} q_i \right) - Q \right]^+ \quad (2.9)$$

In Equation (??) each unit of excess demand is penalized by a factor w. This excess penalty w is decreased by multiplication with a factor ν after ϕ consecutive feasible moves. Similarly, w is increased (multiplied by factor ν^{-1}) after ϕ infeasible iterations.

Every σ iterations, we consider the best found solution (S, v_f) and evaluate the solution for speeds within the interval $[v_f - \tau, v_f + \tau]$. The speed that minimizes $F_2(S, \cdot\,; \alpha, \beta, \gamma)$ is chosen as the new v_f. If a new best feasible solution is found then v_f is kept for another $\sigma\varphi$ iterations. The algorithm is run for I_{max} iteration. However, if a best solution is found and the remaining number of iterations is less than ψ, the remaining number of iterations is updated to 2ψ.

2.4. Experimental settings

We conduct two types of experiments. Section ?? describes our experiments with a single speed profile. The other experiments include two speed profiles and are presented in Section ??. We experiment with sets from ?. The number of customers in these sets ranges between 32 and 80. In this benchmark set, customer locations and their demands are randomly generated by a uniform distribution. To achieve more realistic travel times, the coordinates of these sets were multiplied by 4.9. Note that a test set named "32 k5" means 32 customers including depot given 5 vehicles. Optimal VRP solutions for these sets are available from ?.

2.4.1 Single speed profile

We set up the speed profile with two congested periods, while throughout the rest of the day the speed is set to the free flow speed (Figure ??, left). Many European roads face such a morning and afternoon congestion period (see ?). The right side of Figure ?? depicts the travel time function for each starting time for an arbitrary distance. We set v_2 and v_4 to congestion speed v_c. Furthermore, $v_5 = v_3 = v_1$ are considered as free flow speeds, i.e., they correspond to the decision variable v_f in $F(S, v_f; \alpha, \beta, \gamma)$. The points a, b, c, d, e correspond to 6:00, 9:00, 16:00, 19:00, and 0:00. All links are subjected to this profile. Based on empirical data, the congestion speed is set to 50 km/hr (**?**).

Observe that a starting time after 9:00 is superior to starting time before 9:00, since the latter risks going twice into congestion. Consequently, we assume that all vehicles leave the depot at 9:00. For computing the upper bound, we set a in Equation (**??**) to $c - b$. Considering $v_c = 50$ together with this profile and based on Proposition **??**, leads to the conclusion that $71 \leq v_f^* \leq 91$. Furthermore, by Lemma **??**, the speed for computing the upper bound on $E(S, v_f)$ is $v' = 74$ km/hr.

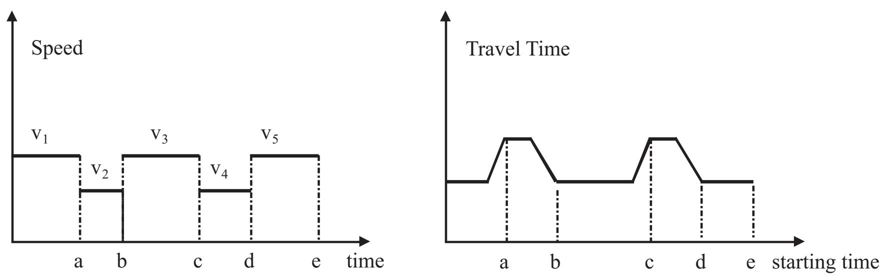

Figure 2.5 Speed and Travel time profile for the experimental setting

The parameters chosen for the tabu search algorithm in section **??** are as follows. For an instance with N nodes, $\eta = \lceil 0.4N \rceil$ closest customers are considered for the 1-interchanges. The tabu tenure length ℓ is randomly chosen in the interval $[15, 25]$. The penalty for infeasible solutions w is set to the initial objective function value divided by 30. Furthermore, $\phi = 5$ and $\nu = \frac{3}{4}$. The maximum number of iterations I_{max} is set to 640 and ψ is set to 160. The parameters relating to v_f are $\sigma = 120$, $\varphi = 1.5$, $\tau = 1$ and the initial speed v_0 is set to 80 km/hr. Unless mentioned otherwise, the cost parameters in $F(S_0, v_f; \alpha, \beta, \gamma)$ are set to: $\alpha = 20$€/hr, $\beta = 1.2$€/liter (diesel), and $\gamma = 11$€/ ton. The latter is based on the actual carbon market cost of the first half of April 2009 **?**. In the computations γ is converted into € per kg, and CO_2 emissions are converted into kg. All runs are performed on a Intel Core Duo with 2.4 GHz and 2 GB of RAM.

Validating the proposed model

To validate the model, we compare the results of $F(S, v_f; 0, 0, 1)$ with the upper bound (UB) and the lower bound (LB) presented in section **??**. Table **??** shows the results for the reduced model where we focus on emissions, i.e., objective function

$F(S, v_f; 0, 0, 1) = E(S, v_f)$. Table **??** gives the CO_2 emissions (kg), the optimal speed v_f, the resulting travel time is $TT(F(S, v_f; 0, 0, 1))$; the last column gives the run times. There is a relation between the amount of CO_2 emissions and the travel time. For example, set 33k5 achieves the lowest emissions and the lowest travel time. The speeds achieving lowest emissions range between 72 and 77 km/hr.

We tested the performance of our model on the standard time-independent VRP. The algorithm reached an average optimality gap of 4% over the different sets. Furthermore, we evaluated the performance of setting the free flow speed to 71 km/hr ($F(S, 71; 0, 0, 1)$), which is the speed which minimized the CO_2 emissions function as observed in Figure **??**. The results showed that adopting this speed leads to an average increase of 2.9 % and 6.0% in emissions and travel times, when compared to the results in Table **??**.

Instance	$F(S, v_f; 0, 0, 1)$	v_f	$TT(F(S, v_f; 0, 0, 1))$	Run time(min)
32 k5	2976	72	3420	4.8
33 k5	2489	74	2836	3.3
33 k6	2777	73	3168	4.2
34 k5	2929	73	3377	3.8
36 k5	3088	75	3449	2.3
37 k5	2516	72	2911	8.5
37 k6	3613	74	4122	4.9
38 k5	2819	73	3209	5.0
39 k5	3175	72	3852	6.6
39 k6	3151	73	3603	7.7
44 k6	3575	72	4122	8.3
45 k6	3644	74	4162	3.0
45 k7	4466	74	4918	2.7
46 k7	3458	74	3912	4.6
48 k7	4156	73	4727	7.4
53 k7	3875	72	4500	13.6
54 k7	4563	77	5023	4.0
55 k9	4174	74	4713	4.0
60 k9	5362	72	6230	16.0
61 k9	4060	74	4536	8.4
62 k8	5104	73	5871	13.2
63 k9	6209	73	7130	13.3
63 k10	5115	73	5869	12.2
64 k9	5379	73	6168	13.6
65 k9	4666	75	5228	6.9
69 k9	4508	75	4996	31.1
80 k10	7062	73	8097	30.9
Average	4034	73	4598	9.0

Table 2.1 Results for the $F(S, v_f; 0, 0, 1)$ model

We generate the bounds using the optimal VRP solutions from **?** for the Augerat sets. Table **??** compares the amount of CO_2 emissions produced using the $F(S, v_f; 0, 0, 1) = E(S, v_f)$ objective function with the bounds. $F(S, v_f; 0, 0, 1)$/LB compares the emissions produced by the E-TDVRP with the lower bound. On average the E-TDRVP is within 4.3% from the lower bound. We note that the LB is computed under the assumption of no congestion. The last column in Table **??** quantifies the gap between UB and LB. The quality of the bounds is very good as the average gap is only 4.7 %. Based on this analysis, we conclude that the solution method is efficient and that the bounds are useful in real-life decision-making.

Set	UB	LB	$F(S, v_f; 0, 0, 1)$	$F(S, v_f; 0, 0, 1)$/LB	UB/LB
32 k5	3081	2904	2976	102.5%	106.1%
33 k5	2546	2443	2489	101.9%	104.2%
33 k6	2838	2738	2777	101.4%	103.6%
34 k5	3016	2880	2929	101.7%	104.7%
36 k5	3135	2957	3088	104.4%	106.0%
37 k5	2602	2479	2516	101.5%	105.0%
37 k6	3684	3510	3613	102.9%	105.0%
38 k5	2822	2706	2819	104.2%	104.3%
39 k5	3220	3056	3175	103.9%	105.4%
39 k6	3240	3071	3151	102.6%	105.5%
44 k6	3634	3463	3575	103.2%	104.9%
45 k6	3648	3483	3644	104.6%	104.7%
45 k7	4458	4229	4466	105.6%	105.4%
46 k7	3531	3386	3458	102.1%	104.3%
48 k7	4172	3960	4156	104.9%	105.3%
53 k7	3897	3735	3875	103.7%	104.3%
54 k7	4553	4320	4563	105.6%	105.4%
55 k9	4078	3961	4174	105.4%	103.0%
60 k9	5238	4998	5362	107.3%	104.8%
61 k9	3925	3830	4060	106.0%	102.5%
62 k8	5041	4771	5104	107.0%	105.7%
63 k9	6344	5980	6209	103.8%	106.1%
63 k10	5042	4843	5115	105.6%	104.1%
64 k9	5437	5164	5379	104.2%	105.3%
65 k9	4498	4356	4666	107.1%	103.2%
69 k9	4438	4298	4508	104.9%	103.3%
80 k10	6900	6512	7062	108.5%	106.0%
Average				104.3%	104.7%

Table 2.2 Upper and Lower bounds for $F(S, v_f; 0, 0, 1)$ (One speed profile).

Numerical results for $F(S, v_f; \alpha, \beta, \gamma)$

Table **??** gives an overview of the cost based solutions for all sets ($\alpha = 20$€/hr,$\beta = 1.2$€/liter and $\gamma = 11$€/ ton). The last three columns give the allocation of costs between its three major components.

				Cost partition		
Set	$F(S_0, v_f; \alpha, \beta, \gamma)$	v_f	CO_2 (kg)	Driver cost	Fuel cost	CO_2 cost
32 k5	2411	84	3092	41.6%	57.0%	1.4%
33 k5	1999	87	2615	40.4%	58.1%	1.4%
33 k6	2267	85	2924	41.3%	57.3%	1.4%
34 k5	2380	86	3061	41.4%	57.2%	1.4%
36 k5	2494	85	3221	41.2%	57.4%	1.4%
37 k5	2043	85	2636	41.2%	57.3%	1.4%
37 k6	2934	86	3799	41.0%	57.6%	1.4%
38 k5	2247	84	2937	40.5%	58.1%	1.4%
39 k5	2671	85	3423	41.6%	56.9%	1.4%
39 k6	2788	86	3626	40.8%	57.8%	1.4%
44 k6	2952	87	3893	39.9%	58.6%	1.5%
45 k6	2984	82	3750	42.8%	55.9%	1.4%
45 k7	3837	87	4958	41.1%	57.4%	1.4%
46 k7	2856	84	3680	41.3%	57.3%	1.4%
48 k7	3341	86	4349	40.7%	57.8%	1.4%
53 k7	3245	86	4208	40.9%	57.6%	1.4%
54 k7	3714	82	4676	42.7%	55.9%	1.4%
55 k9	3463	80	4327	43.1%	55.5%	1.4%
60 k9	4534	80	5632	43.4%	55.2%	1.4%
61 k9	3341	82	4247	42.1%	56.5%	1.4%
62 k8	4370	86	5655	41.1%	57.5%	1.4%
63 k9	5026	86	6476	41.3%	57.3%	1.4%
63 k10	4414	80	5505	43.2%	55.4%	1.4%
64 k9	4612	85	5916	41.6%	57.0%	1.4%
65 k9	3604	83	4628	41.5%	57.1%	1.4%
69 k9	3679	84	4773	40.9%	57.7%	1.4%
80 k10	5631	84	7187	41.9%	56.7%	1.4%
Average				41.5%	57.1%	1.4%

Table 2.3 Results for $F(S_0, v_f; \alpha, \beta, \gamma)$ (One speed profile)

The vast majority of costs are attributed to driving cost and fuel cost, which together on average account for 98.6 % of the costs. The cost of CO_2 is rather insignificant when compared to others: 1.4% on average. Specifically, the current market cost of CO_2 has no tangible economic impact on the results. However, due to the direct relation between fuel consumption and CO_2, speeds associated with this model are between 80 and 87 km/hr. In fact, in all instances the resulting speed is strictly less than 90 km/hr. This implies that even with the current cost structure speeds are likely to be less than 90 km/hr.

The reduced models: $F(S, v_f; 0, 0, 1)$ **and** $F(S, v_f; 1, 0, 0)$

We analyze the results for the two reduced models. The first one optimizes only on the emissions ($F(S, v_f; 0, 0, 1)$), the second one focuses on travel times ($F(S, v_f; 1, 0, 0)$). This facilitates a trade-off analysis between travel times and emissions. We ran the settings $F(S, 80; 1, 0, 0)$ and $F(S, 90; 1, 0, 0)$, i.e., we fixed the speed limits to 80 and 90 respectively. The truck speed limits in most European countries is 80 km/hr or 90 km/hr so these two specific settings depict an interesting European benchmark.

Below, we define the resulting travel times from the (reduced) model with objective function $F(S, v_f; 0, 0, 1)$ as $TT(F(S, v_f; 0, 0, 1))$. Specifically, the objective function only considers emissions in its optimization $F(S, v_f; 0, 0, 1)$. The resulting solution (S, v_f) is then evaluated in terms of their travel times ($TT(F(S, v_f; 0, 0, 1))$). Similarly, we define the emissions corresponding to the (reduced) model with objective function $F(S, 80; 1, 0, 0)$ as $E(F(S, 80; 1, 0, 0))$. This means that we generate a solution (S, v_f) by minimizing the travel time and then evaluate this solution in terms of its emissions $E(F(S, 80; 1, 0, 0))$. We use the following notation.

$$
\begin{aligned}
\Delta TT_{80} &= \frac{F(S, 80; 1, 0, 0) - TT(F(S, v_f; 0, 0, 1))}{F(S, 80; 1, 0, 0)} \\
\Delta E_{80} &= \frac{E(F(S, 80; 0, 0, 1)) - F(S, v_f; 0, 0, 1)}{E(F(S, 80; 1, 0, 0))} \\
\Delta TT_{90} &= \frac{F(S, 90; 1, 0, 0) - TT(F(S, v_f; 0, 0, 1))}{F(S, 90; 1, 0, 0)} \\
\Delta E_{90} &= \frac{E(F(S, 90; 0, 0, 1)) - F(S, v_f; 0, 0, 1)}{E(F(S, 90; 1, 0, 0))}
\end{aligned}
$$

Instance	ΔE_{80}	ΔE_{90}	ΔTT_{80}	ΔTT_{90}
32 k5	0.8%	9.2%	-10.3%	-20.2%
33 k5	4.2%	8.5%	-4.9%	-20.6%
33 k6	5.3%	12.4%	-4.0%	-15.1%
34 k5	2.8%	12.6%	-7.3%	-19.1%
36 k5	1.4%	8.2%	-6.5%	-20.2%
37 k5	2.1%	9.0%	-9.0%	-21.5%
37 k6	1.1%	11.2%	-7.3%	-17.5%
38 k5	1.4%	16.1%	-11.1%	-17.0%
39 k5	8.5%	12.0%	-6.4%	-21.0%
39 k6	2.0%	16.0%	-8.9%	-10.3%
44 k6	5.3%	9.5%	-5.6%	-20.4%
45 k6	3.9%	9.7%	-7.6%	-19.2%
45 k7	0.8%	9.3%	-4.5%	-12.0%
46 k7	1.4%	14.3%	-8.1%	-13.2%
48 k7	2.1%	12.3%	-6.3%	-13.5%
53 k7	5.1%	9.9%	-5.8%	-20.5%
54 k7	1.5%	6.7%	-5.1%	-17.1%
55 k9	2.1%	12.7%	-7.4%	-18.0%
60 k9	4.9%	11.9%	-5.5%	-17.0%
61 k9	2.4%	10.3%	-7.7%	-22.6%
62 k8	6.4%	13.9%	-2.5%	-11.5%
63 k9	5.3%	8.6%	-3.6%	-17.6%
63 k10	3.4%	11.7%	-6.6%	-16.7%
64 k9	4.7%	14.2%	-4.4%	-13.1%
65 k9	3.3%	12.0%	-5.0%	-15.3%
69 k9	2.5%	10.2%	-6.2%	-20.3%
80 k10	0.7%	14.5%	-9.7%	-10.5%
Average	3.2%	11.4%	-6.6%	-17.1%

Table 2.4 Trade-off between emissions and travel times (One speed profile)

Table **??** gives an overview of the potential savings in emissions, compared to the amount of travel time needed to be sacrificed in order to achieve this saving. Columns 2 and 3 exhibit the increase in travel times, while the last two columns exhibit the decrease in emissions. Considering the speed limit of 90 km/hr, Table **??** implies that an average reduction of 11.4% in CO_2 emissions can be achieved. However, such a reduction will result in an average increase in travel time by 17.1 %. Looking to the situation with a speed limit of 80 km/hr an average of 3.2% reduction in emissions can be achieved, which would require an average increase of 6.6% in travel time. These results show that in the case of 90 km/hr a substantial reduction of CO_2 emissions can be achieved.

2.4.2 Two speed profiles

In this section we consider two speed profiles. In addition to the profile considered in section **??**, we add a similar profile but with a congestion speed equal to 70 km/hr **?**), instead of 50 km/hr. The two profiles were randomly assigned to the various distances in the problems.

We ran $F(S_0, v_f; \alpha, \beta, \gamma)$ for the two speed profile setting. Table **??** gives an overview of the performance of the solutions for all sets, the last three columns give the allocation of cost between its three components. Similar to the single speed profile case, the majority of costs are attributed to driving cost and fuel cost. However, due to the direct relation between fuel consumption and CO_2, speeds associated with this model are between 81 and 85 km/hr.

Both for the single and for the two speed profile case the resulting speeds (v_f) are within the same range. For the single speed profile case, where congestion speed is 50 km/hr, the best strategy to limit fuel and emission costs is to try to avoid congestion, and thus speeds are higher than $v^*(= 71$ km/hr). For the two speed profile case, the congestion speeds are 70 km/hr and 50 km/hr. The congestion speed of 70 km/hr is rather similar to the v^*. For these cases, minimizing fuel and emission costs leads to prefer congestion and to reduce the speed. However, this is countered by the travel time cost.

As the resulting speeds for both one and two speed profiles are within similar ranges, we examine the case where we set v_f to 85 km/hr for all sets. This value is observed in 14 sets from Table **??**. Furthermore, one value for all sets might be easier from an operational point of view. Similar to the above definitions, we define the following relations.

$$
\begin{aligned}
\Delta(TT)_{85} &= \frac{TT(F(S,85;\alpha,\beta,\gamma)) - TT(F(S,v_f;\alpha,\beta,\gamma))}{TT(F(S,85;\alpha,\beta,\gamma))} \\
\Delta(E)_{85} &= \frac{E(F(S,85;\alpha,\beta,\gamma)) - E(F(S,v_f;\alpha,\beta,\gamma))}{E(F(S,85;\alpha,\beta,\gamma))} \\
\Delta(F)_{85} &= \frac{F(S,85;\alpha,\beta,\gamma) - F(S,v_f;\alpha,\beta,\gamma)}{F(S,85;\alpha,\beta,\gamma)} \\
\Lambda(F)_{85} &= F(S,85;\alpha,\beta,\gamma) - F(S,v_f;\alpha,\beta,\gamma)
\end{aligned}
$$

Set	$F(S_0, v_f; \alpha, \beta, \gamma)$	v_f	CO_2 (kg)	Cost partition		
				Driver cost	Fuel cost	CO_2 cost
32 k5	2487	85	3231	40.8%	57.7%	1.4%
33 k5	1968	85	2564	40.7%	57.9%	1.4%
33 k6	2234	83	2887	41.1%	57.4%	1.4%
34 k5	2332	85	3039	40.7%	57.9%	1.4%
36 k5	2450	84	3161	41.2%	57.3%	1.4%
37 k5	2004	85	2625	40.4%	58.2%	1.4%
37 k6	2925	85	3823	40.5%	58.1%	1.4%
38 k5	2199	85	2874	40.5%	58.1%	1.4%
39 k5	2632	85	2874	40.8%	57.7%	1.4%
39 k6	2763	84	3571	41.1%	57.4%	1.4%
44 k6	2875	84	3728	40.9%	57.6%	1.4%
45 k6	2850	85	3729	40.4%	58.2%	1.4%
45 k7	3633	85	4744	40.5%	58.0%	1.4%
46 k7	2781	85	3641	40.4%	58.2%	1.4%
48 k7	3393	83	4371	41.3%	57.2%	1.4%
53 k7	3221	85	4198	40.6%	57.9%	1.4%
54 k7	3736	84	4851	40.9%	57.7%	1.4%
55 k9	3319	81	4243	41.8%	56.8%	1.4%
60 k9	4371	82	5593	41.7%	56.9%	1.4%
61 k9	3189	82	4114	41.2%	57.3%	1.4%
62 k8	4279	85	5588	40.5%	58.0%	1.4%
63 k9	4976	83	6394	41.5%	57.1%	1.4%
63 k10	4069	84	5300	40.7%	57.9%	1.4%
64 k9	4554	85	5951	40.5%	58.1%	1.4%
65 k9	3826	85	5027	40.2%	58.4%	1.4%
69 k9	3639	85	4804	39.9%	58.7%	1.5%
80 k10	5673	82	7250	41.8%	56.8%	1.4%
average				40.8%	57.7%	1.4%

Table 2.5 Results for $F(S, v_f; \alpha, \beta, \gamma)$ (Two speed profiles)

Table **??** exhibits the results for setting v_f to 85 km/hr. We note that the differences are not substantial. The cost increases only by 1.1 % on average. However, in absolute values, as depicted by $\Lambda(F)_{85}$, differences of up to 144.6€ are observed. For the analyzed sets, we conclude that setting the vehicle speed to 85 km/hr results in a good compromise between travel time minimization and emissions minimization.

Set	$F(S_0, 85; \alpha, \beta, \gamma)$	$\Delta(TT)_{85}$	$\Delta(E)_{85}$	$\Delta(F)_{85}$	$\Lambda(F)_{85}$
32 k5	2489	0.0%	0%	0%	0
33 k5	1968	0.0%	0%	0%	0.0
33 k6	2234	2.9%	5%	4%	100.1
34 k5	2384	0.0%	0%	0%	0.0
36 k5	2450	0.7%	5%	4%	89.0
37 k5	2040	0.0%	0%	0%	0.0
37 k6	2960	0.0%	0%	0%	0.0
38 k5	2229	0.0%	0%	0%	0.0
39 k5	2632	0.9%	2%	2%	47.9
39 k6	2763	1.7%	2%	2%	54.8
44 k6	2875	3.8%	5%	4%	129.0
45 k6	2899	0.0%	0%	0%	0.0
45 k7	3643	0.0%	0%	0%	0.0
46 k7	2908	0.0%	0%	0%	0.0
48 k7	3393	-2.0%	2%	0%	8.5
53 k7	3258	0.0%	0%	0%	0.0
54 k7	3736	-0.1%	1%	1%	25.9
55 k9	3319	-3.9%	4%	1%	25.4
60 k9	4371	-1.6%	2%	1%	35.8
61 k9	3189	-0.2%	5%	3%	104.7
62 k8	4282	0.0%	0%	0%	0.0
63 k9	4976	-0.5%	3%	2%	78.6
63 k10	4069	2.9%	4%	4%	144.6
64 k9	4738	0.0%	0%	0%	0.0
65 k9	3891	0.0%	0%	0%	0.0
69 k9	3665	0.0%	0%	0%	0.0
80 k10	5673	-0.9%	4%	2%	124.5
average		0.1%	1.7%	1.1%	35.9

Table 2.6 Two speed profiles: Comparison of $F(S, v_f; \alpha, \beta, \gamma)$ and $F(S, 85; \alpha, \beta, \gamma)$

2.5. Conclusions and future research

Carrier companies need to consider employing greener practices in the future. This chapter establishes a framework for modeling CO_2 emissions in a time-dependent VRP context (E-TDVRP). It proposes an efficient solution method for the E-TDVRP. We showed the potential reduction in emissions when weighed against travel time.

Furthermore, reducing emissions has a positive impact with respect to reducing cost.

We analyzed the effect of limiting vehicle speed in a time-dependent VRP setting on the generated amount of CO_2 emissions. Time-dependency was established by subjecting distances to variable speed profiles. CO_2 emissions were modeled as a function of speed. The speed which optimizes CO_2 emissions per km was not observed in any of our experimental settings. This is explained by the relation between limiting free flow speed and the likelihood of running into congestion. In congestion, the vehicle is forced to drive slower and thus emitting more CO_2. Thus, increasing the free flow speed decreases the amount of time spent in congestion, and by that results in less emissions.

Based on the optimal VRP solutions in the literature, we constructed an upper and lower bound for the CO_2 emissions. We showed that these bounds are tight, on average the gap was 4.7%. From a practical point of view these bounds are simple and fast to calculate. They require standard VRP solutions commonly used in practice. Thus, they enable assessing the potential reduction in CO_2 emissions.

Considering CO_2 emissions, the results of the E-TDVRP were rather close to the lower bound, with a difference of 4.3% on average. This indicates that the chosen solution strategy was adequate. We compared two extreme cases of the E-TDVRP model, considering only travel times and considering only CO_2 emissions. For the first case, two possible speed limits were considered, 80 and 90 km/hr. For a speed limit of 90 km/hr our results showed that achieving an average reduction of 11.4% in CO_2 emissions required increasing travel time by 17.1%. However, considering a speed limit of 80 km/hr, an average increase of 6.6% in travel achieves a reduction of 3.2% in CO_2 emissions.

Considering one and two congestion speeds, the experiments showed that from a realistic cost perspective it is beneficial for vehicles to go below 90 km/hr. We also showed that adopting a speed limit of 85 km/hr yields good results in terms of total cost over all sets. Finally, we conclude that minimizing CO_2 emissions by limiting vehicle speed can be costly in terms of travel time. However, limiting vehicle speed to a certain extent might be both cost and emission effective.

Further research can be conducted in more accurate emissions modeling. For example, the weight of the vehicle has an impact on the amount of emissions, and in the VRP context this can be incorporated as function of satisfied demand. Additionally, more sophisticated travel time profiles can be constructed to encompass acceleration and deceleration, which in return will enable more accurate emissions modeling. Furthermore, the proposed model can be employed to investigate the costs and benefits of using alternative fuels in the context of VRP.

We have shown that the current market cost of CO_2 has no tangible economic impact on the results. However, exploring this value is yet another possible extension to this work. Examining the impact of this value on transportation related strategies is a relevant research question.

Chapter 3

TDVRP Subject to Time Delay Perturbations

3.1. Introduction

Following the operations of a number of carrier companies, we observed that during the execution phase of their vehicle routing schedules many unexpected delays were encountered. Most of these delays were caused by customers' distribution centers not being ready to receive their goods. Equivalently, such delays could also occur at the carrier's depot when the truck is not fully loaded at the scheduled starting time. When confronted with such unanticipated delays, the trucks are forced to wait until the customer is ready. This is mainly because the trucks were filled to serve a certain subset of clients in a particular order, and hence altering a truck's schedule will imply going back to the depot and/or rearranging the filling order, which is costly from an operational standpoint. Consequently, the carrier companies did not want to update their schedules on-line in compliance with these delays, but would rather have a schedule that is better protected against this specific type of time delays.

The above described situation is translated into a specific vehicle routing setting (VRP), taking into account time-dependent information with respect to travel times. In essence, this means that the travel time between two consecutive customers not only depends on the distance, but also on the starting time. In this time-dependent scheduling setting, unanticipated delays at a customer will influence the starting time of going to the subsequent customers and thus on the expected travel time between these customers. Consequently, assessing the impact of these delays is less straightforward than in a time-independent environment (with constant speeds or travel times).

This chapter focuses on Time-Dependent Vehicle Routing Problems (TDVRP). We model traffic by discrete speed zones, where speeds are derived from real-life data and translated into travel time functions that satisfy the FIFO principle (see also **?**). In addition to this, we build a TDVRP model that optimizes the routing costs subject to unanticipated delays which are modeled as perturbations. We refer to this model as P-TDVRP. We explore the existence of routing schedules that perform well and we identify the structural properties of these schedules which minimize the resulting weighted cumulative delay. We pay specific attention to the analysis of the properties of the generated routes. **?** argue that an analytical analysis of the vehicle routing problem brings new insight into the algorithmic structure and makes performance analysis of classical algorithms possible.

The main contributions of this chapter are the following:

- This chapter proposes a P-TDVRP approach for dealing with unexpected delays at the various nodes in a TDVRP setting. An optimization procedure is proposed to trace solutions that perform well under perturbations. We show that these solutions in many cases differ substantially from the ones obtained by the TDVRP (i. e., not taking into account these delays). The optimization technique proposed is, to our knowledge, the first application of the methodology proposed by **?** to a discrete problem coupled with a Tabu Search heuristic. The method evaluates disruptions and considers their average value as the objective of the optimization.

- We identify situations capable of absorbing delays, i. e., where inserting a delay will lead to an increase in travel time that is less than or equal to the expected delay length itself. This is namely observed when there is a speed increase, since a delay at a customer might imply starting at a zone with higher speeds. We evaluate the cost-benefit trade-off when using P-TDVRP routing schedules. We weigh the costs and gains of the P-TDVRP routing schedules compared to the TDVRP ones. In the majority of the tested cases, the results show higher gains than costs.

- The specific characteristics and behavior of the P-TDVRP routing schedules are identified. We observe that the P-TDVRP routing schedules have higher absorption potential and are comprised of less disperse routes, i. e., routes where tours are more evenly lengthened and more links are likely to be in the vicinity of absorption zones. These properties are formulated into proposition that are observed against the benchmark of the TDVRP routing schedules obtained.

This chapter is organized as follows. In Section ??, we present a literature survey. In Section ??, we present our time-dependent VRP model, and in Section ??, we define the Perturbed TDVRP model (i. e. P-TDVRP). In Section ??, we present an analysis of solution structures that are likely to perform better under perturbations. Our chosen experimental setting is presented in Section ??. In Section ??, we formulate propositions, with respect to the solution structure, and show the results. Finally in Section ??, we highlight the main findings and discuss directions for future research.

3.2. Literature review

Much research has been conducted on Vehicle Routing Problems (VRP). For a comprehensive overview, the reader is referred to ??. The vast majority of research focuses on a constant speed environment, i. e., where speeds remain constant throughout the day. Realizing that speeds and their associated travel times depend on the time of day, more and more research is dealing with VRPs in a time-dependent setting (TDVRP).

For the literature on time-dependent VRP the reader is referred to Section ??. We mention that we model time dependency similar to ?, ?, and ?. Thus, our model adheres to the FIFO assumption.

Over the past years, optimizing under uncertainty has attracted much research. This is driven by the realization that in a modeling process parameters are assessed and are inserted in the given constraints. However, in many cases a great deal of uncertainty lies in their assessments. In principle this means that, while sufficiently good solutions are found for a specific parameter set, their applicability is limited, since they do not take into account parameter uncertainty. Robust optimization as presented by ??, is essentially defined over a convex space. ? applied this framework to a VRP with uncertain demand on a set of fixed customers. While this area has evolved throughout the past years, it would be difficult to fit our problem into it, since disruptions occur in a time-dependent environment and their impact on the costs is not straightforward. Another view on robustness is given in ?, who address uncertainty in travel time in TSP by applying a robust deviation criterion.

? defines robustness as the insensitivity of a solution with respect to changes in the environment in which this solution is implemented. He further notes that while there are a number of powerful tools to obtain robust solutions, they are very difficult to implement with local search techniques. Based on the work of ?, he proposes an approach that incorporates perturbations on given solutions which are then evaluated through the objective function. This approach was presented in the context of robust continuous optimization using local search. We choose to adapt this approach to our discrete TDVRP setting (due to its simplicity, relevance, and adaptability to Tabu Search). We show that significant gains can be obtained by using solutions produced under perturbations.

3.3. Time-Dependent Vehicle Routing Problems

Formally, the vehicle routing problem can be represented by a complete directed graph $G = (V, A)$ where $V = \{0, 1, \ldots, n\}$ is a set of vertices and $A = \{(i, j) : i <> j \in V\}$ is the set of directed links. The vertex 0 denotes the depot; the other vertices of V represent customers, the number of available vehicles is N. For each customer, a non-negative demand d_i is given where $d_0 = 0$.

The objective is to find, for the given number of vehicles N, the minimum costs where the following conditions hold: every customer is visited exactly once by exactly one vehicle, all vehicle routes start and end at the single depot, every route has a total demand not exceeding the vehicle capacity Q.

We define a solution as a set S with s routes $\{R_1, R_2, ..., R_s\}$, where $s \leq N$, $R_r = (0, .., i, ..., 0)$, i. e., each route R_r begins and ends at the depot. We write $i \in R_r$ if the vertex $i \geq 1$ is part of the route R_r, where each vertex belongs to exactly one route. We write $(i, j) \in R_r$ if i and j are two consecutive vertices in R_r.

Let $t(r) \in \{t_1, ..., t_k\}$ denote the starting time of route R_r. Let $t(r, i)$ denote the departure time at node i, for route R_r. $TT_{i,j,t(r,i)}$ represents the travel time between node i and node j when starting at time $t(r, i)$. Furthermore, we define a shift length w. Any route length exceeding w will result in an overtime penalty of p, for every excess time unit. In Equation (**??**), we denote by $F_{norm}(S)$ the objective function value associated with the solution S.

$$F_{norm}(S) = \sum_{r} \sum_{t(r)} \sum_{(i,j) \in R_r} TT_{i,j,t(i,r)} + p \sum_{r} \left[\sum_{t(r)} \sum_{(i,j) \in R_r} TT_{i,j,t(i,r)} - w \right]^{+} \tag{3.1}$$

In the above equation $[x]^{+} = \max\{0, x\}$; p is a positive parameter. F_{norm} is comprised of total travel and a penalty component on overtime.

The time dependency of travel times throughout the planning horizon is driven by the changing speeds in the different time zones. Since the number of links in a fully connected directed graph is enormous, collecting data for each link and each time zone is infeasible. Instead, we assume that the same speed profile applies to all links. Indeed it seems reasonable to assume that most motor ways, on average, follow the same pattern of having a morning and evening congestion period (see also ? for a similar reasoning). The FIFO assumption, also known as the non-surpassing assumption, infers that a vehicle starting to traverse a link at any moment t will arrive earlier than other vehicles starting to traverse the same link after time t. This is valid in the VRP case in particular, since one assumes that the vehicles are identical in terms of performances, and thus vehicles cannot surpass each other. As in Chapter 2, we adopt this assumption.

In essence, throughout the time horizon we construct a speed profile, which is translated into a travel time function satisfying the FIFO assumption. The generation of a travel time profile can be viewed as an integration over the speed profile. An example for a given distance d is illustrated in Figure ??. The right side of this figure depicts the travel time pattern associated with a single step speed increase as a function of the starting time. The main intuition for the profile is that during the first period (up to $a - TT_a$), it takes TT_a time units to traverse the link, since throughout that period the vehicle will be driving at a speed v_a along the entire link; however, for starting time $a - TT_a$ up to point a (the transient zone), the vehicle will start traversing part of the link with speed v_a and the remainder with a speed v_b. Starting at point a the speed will remain constant at v_b, resulting into travel time TT_b. Thus the travel time is a continuous function over the starting times.

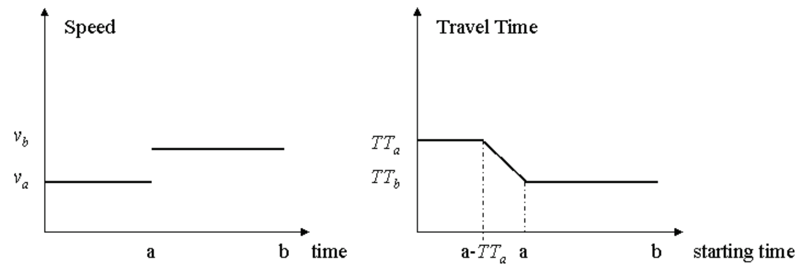

Figure 3.1 The conversion of speed into travel times

The start of the transient zone occurs at point $a - TT_a$. The linearity in the function stems from the stepwise speed change which imposes different speeds for continuous time intervals. For distance d to be traversed, the slope in the transient zone is independent of the distance, since the slope can be defined as $\frac{TT_b - TT_a}{TT_a}$. By substituting TT_a with $\frac{d}{v_a}$ and TT_b with $\frac{d}{v_b}$, we obtain that the slope is equal to $\frac{v_a - v_b}{v_b}$, while the intersection with the Travel Time axis, is a function of the distance and is equal to $\frac{(v_b - v_a)}{v_b} a + \frac{d}{v_b}$. Thus, the specific travel time functions differ from one link to another, since the distances between the customers are different. We define $g(t)$ as the travel time function associated with any starting time t for Figure **??**.

$$g(t) = \begin{cases} TT_a & \text{if } t \leq a - TT_a \\ \frac{v_a - v_b}{v_b} t + \frac{(v_b - v_a)}{v_b} a + \frac{d}{v_b} & \text{if } a - TT_a < t < a \\ TT_b & \text{if } t \geq a \end{cases}$$

For starting times within $(0, a - TT_a)$ FIFO is satisfied since the speed associated with these starting times is constant, given two starting times within this interval t and $t + \Delta$, both will arrive after TT_a time units from their departure. By a similar reasoning, FIFO holds for starting times after a. The line in the transient zone $(a - TT_a, a)$ can be viewed as a consequence of the FIFO assumption as well. Given two starting times t and $t+\Delta$, both in the transient zone, the arrival times are $t+g(t)$ and $t + \Delta + g(t + \Delta)$ respectively, the difference between the arrival times of $t + \Delta$ and t is $\Delta + \frac{v_a - v_b}{v_b}\Delta > 0$. And thus the FIFO assumption holds in the transient zone. While this was shown for a speed increase, similar functions can be constructed for speed drops as well. Note that this approach to modeling travel times is actually equivalent to integrating the distance over the different speeds.

3.4. Perturbed Time-Dependent Vehicle Routing Problems

We model disruptions in the above formulated Time-Dependent Vehicle Routing Problem by time perturbations. Optimizing problems while taking into account uncertainty, with respect to the actual performance of a given solution, is a common problem. In the genetic algorithms literature, **?**, this has been dealt with by perturbing the solution before evaluating it, i. e., by replacing the objective function $f(x)$ with $F(x) = f(x + \delta)$ where δ is usually a normally distributed noise. The basic idea behind this approach is that the mechanism will prefer solutions on wide hills as apposed to narrow peaks. A solution is regarded good if its expected performance with respect to a number of randomly generated perturbations is good. **?** further implemented the mechanism in a Tabu Search setting for continuous functions.

We adopted this methodology to the TDVRP by modeling disruptions as perturbations, leading to the P-TDVRP. We assume that these disruptions can occur at any customer location. We also assume that only one disruption can occur during the day. Given a solution comprised of m links, we generate m scenarios, corresponding to a single perturbation per link. We evaluate the average performance of a solution over all scenarios and consider this average value as the target function value for the solution.

For each m, the model assumes one perturbation. This approach is motivated from a project-and-machine scheduling context. More specifically, e. g., **?**, **?**, and **??** focused on the effect of a single disruption of each of the individual jobs, rather than on all possible disruption interactions. This is due to the fact that the probability of multiple disruptions occurring simultaneously is far lower than that of a single disruption occurrence. Moreover a solution that performs well under a single disruption is likely to perform well under multiple ones as well (**?**). Single disruption assumptions are also used in the context of telecommunications network reliability, see, e. g., **?**.

The P-TDVRP objective function is denoted by $F_{pert}(S)$, where $S(\delta_i)$ depicts the scenario where a perturbation of size δ is placed before the i^{th} link in solution S. $F_{pert}(S)$ is expressed in Equation (**??**).

$$F_{pert}(S) = \frac{1}{m}\sum_{i=1}^{m} F_{norm}(S(\delta_i)) \tag{3.2}$$

By Equation (**??**), in a given solution each link will be perturbed once. Whenever a perturbation occurs, it will increase travel time. Since our model complies with the FIFO assumption, the costs will inevitably rise. In order to compare our results, we define costs (C) and gains (G) by Equations (**??**) and (**??**), respectively. We define $S^* = \arg\min_S F_{norm}(S)$ and $Z^* = \arg\min_S F_{pert}(S)$.

$$C = F_{norm}(Z^*) - F_{norm}(S^*) \tag{3.3}$$

$$G = F_{pert}(S^*) - F_{pert}(Z^*) \tag{3.4}$$

Equation (**??**) depicts the potential loss in the normal target function value when using the value of the perturbed solution, whereas Equation (**??**) describes the potential gain of using the optimal perturbed solution as opposed to the optimal normal solution in the perturbed setting. By definition, both *Costs* and *Gains* are positive. Subtracting the *Costs* from the *Gains* gives the trade-off one needs to consider. A positive value of the trade-off indicates that the potential gain in the perturbed setting prevails over the cost of using the normal solution. Similarly, a negative trade-off indicates that the potential cost of using the solutions from the perturbed optimization is higher than the gains expected from it.

Figure **??** gives an illustration of P-TDVRP and TDVRP solutions. On the categorical X-axis we have two possible solutions S^* and Z^*, where the $F_{norm}(S^*) < F_{norm}(Z^*)$, i. e., S^* outperforms Z^* in the TDVRP optimization, perturbing each of the solutions results in an increase in costs, the average of the perturbed values are given by $F_{pert}(S^*)$ and $F_{pert}(Z^*)$, respectively. Therefore, while S^* is superior to Z^* in the TDVRP, the opposite holds for the P-TDVRP.

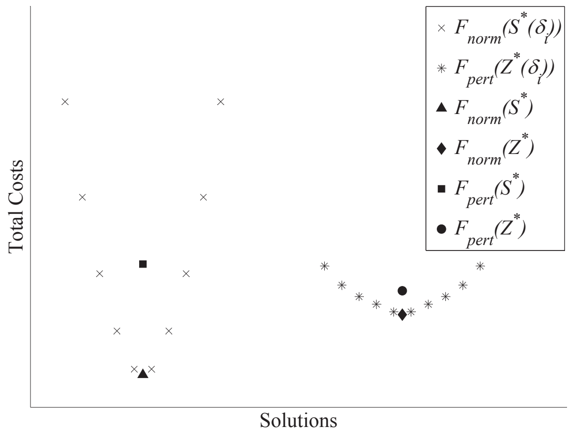

Figure 3.2 Perturbations in solutions

The strong point of this approach is that it does not call for any change to the Tabu Search structure, but rather evaluates the target function in a different manner. The drawback is an additional computation time. We further note that, to the best of our knowledge, this is the first application of the methodology proposed by **?**, to a Tabu Search environment with a discrete problem.

3.5. Analysis of the P-TDVRP: The absorption effect

In this section, we analyze the consequences of the Perturbed Time-Dependent Vehicle Routing model on travel times. More specifically, we investigate the effect of a perturbation on the travel times. We show that the net effect of a time delay δ on the total travel times is not always clear-cut (i. e., the effect is not merely adding the time delay to the original travel times). Indeed, there exist situations in which the increase in the travel times is less than the time delay. We denote this behavior as the absorption effect AE, which can be positive (i. e., relatively less time than the delay is added to the travel times), negative (i. e., relatively more time than the delay is added to the travel times), or zero. Note that the AE is only relevant for the TDVRP;

since in the time-independent VRP the speed is constant throughout the day, adding a perturbation of size δ will constitute a constant addition to the travel time. Thus, in the time-independent VRP setting, adding a perturbation will not influence the optimal solution sequence. In the remainder of this section we first show situations where on average there exists a positive absorption. After identifying situations with positive absorption potential we elaborate the magnitude of the absorption, for some discrete.

3.5.1 The absorption effect defined with one link

Figure **??** gives an illustration of a positive absorption effect. The left side of this figure depicts the perturbation as perceived, while the right side depicts the actual resulting profile. Similar to Figure **??**, the travel time is given as a function of the starting time. The solid lines are the normal travel times, i. e. without a perturbation, while the dashed lines give the perturbed travel time. In the perceived situation the perturbed travel time is an addition of δ to the original one. However since the perturbation is added to the starting time of traversing a link, the actual perturbed travel time differs from the perception, the transient zone is $[a - TT_a - \delta, a - \delta]$. This is due to the fact that when subjected to a perturbation, the starting time $a - TT_a - \delta$ will be delayed by δ and eventually start at $a - TT_a$, and similarly $a - \delta$ will shift to a.

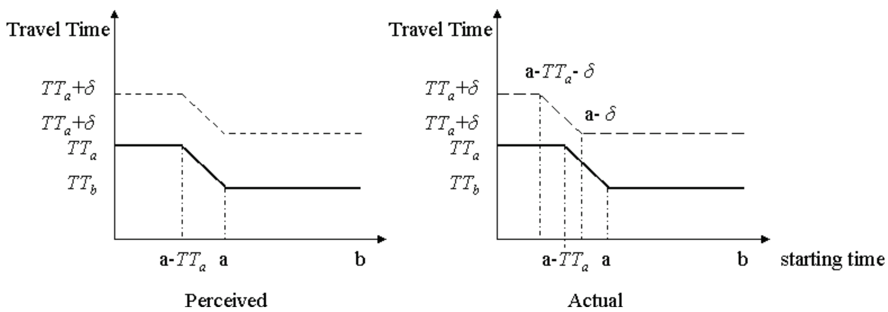

Figure 3.3 Travel time profile with perturbations

For the subsequent analysis, we define the travel time plus a perturbation of size δ as the *Perceived* TT, *i.e* $TT_{ijt} + \delta$. We further define the actual travel time when a perturbation of size δ is actually imposed to the solution as *Actual* TT, i. e. $TT_{ijt,\delta}$. The absorption effect (AE) is then quantified as the difference between *perceived* TT and *Actual TT*. A positive AE implies that adding a perturbation of size δ to an edge that initially had TT_{ijt} resulted in a travel time that is less than $TT_{ijt} + \delta$.

The AE effect is quantified as follows. Assume that the starting time of traversing the link between city i and j is uniformly distributed over the planning horizon $[0, b]$. In the perceived situation, the average resulting perturbed travel time is equal to the following.

$$\begin{aligned} TT_{ijt} + \delta &= \frac{1}{b}\int_0^{a-TT_a} (TT_a + \delta)\, dt \\ &+ \frac{1}{b}\int_{a-TT_a}^{a} \left[\frac{TT_b - TT_a}{TT_a}t + \left(TT_b + a\frac{TT_a - TT_b}{TT_a} + \delta\right)\right] dt \\ &+ \frac{1}{b}\int_a^b (TT_b + \delta)\, dt \\ &= \frac{1}{b}\left(a\,(TT_a - TT_b) - \frac{TT_a^2}{2} + \frac{TT_a TT_b}{2} + bTT_b\right) + \delta \end{aligned} \tag{3.5}$$

However, since the travel time would start to drop at point $(a - TT_a - \delta)$ rather than at time $(a - TT_a)$, and starting point $a - \delta$ the travel time will be $TT_b + \delta$, the actual travel time becomes:

$$\begin{aligned} TT_{ijt,\delta} &= \frac{1}{b}\int_0^{a-TT_a-\delta} (TT_a + \delta)\, dt \\ &+ \frac{1}{b}\int_{a-TT_a-\delta}^{a-\delta} \left[\frac{TT_b - TT_a}{TT_a}t + \left(TT_b + (a-\delta)\frac{TT_a - TT_b}{TT_a} + \delta\right)\right] dt \\ &+ \frac{1}{b}\int_{a-\delta}^{b} (TT_b + \delta)\, dt \\ &= \frac{1}{b}\left(a\,(TT_a - TT_b) - \frac{TT_a^2}{2} - \delta TT_a \right. \\ &+ \left. \frac{TT_a TT_b}{2} + \delta TT_b + b\,(TT_b + \delta)\right) \end{aligned} \tag{3.6}$$

The average absorption effect (AE) is then obtained by subtracting Equation (**??**) from Equation (**??**):

$$\overline{AE} = \frac{\delta\,(TT_a - TT_b)}{b} \tag{3.7}$$

Equation **??** can be explained intuitively. Comparing the actual and perceived travel times in Figure **??**, a total of $\delta(TT_a + \delta)$ is lost, a total of $\delta(TT_b + \delta)$ is gained, and this is averaged over b.

Plotting the perceived travel time along with the actual one, results in a parallelogram where the upper line is the perception and the lower one is the actual. The parallelogram area on the right side of Figure **??** illustrates the average absorption effect $\overline{AE}$ for the case where $TT_a > \delta$. The left side of the figure illustrates $\overline{AE}$ when $TT_a < \delta$. Figure **??** highlights the average absorption over a set of starting times. In the P-TDVRP model, each link will start in a discrete point: if that point is within $[a - TT_a - \delta, a]$, an absorption will take place; however, if the starting point is elsewhere, the absorption will be zero.

We note that a positive absorption results on average when assuming a distribution on the starting times. However, in a discrete schedule, links are traversed in a sequence by which each link has a unique starting point which dictates the starting points of its successive links.

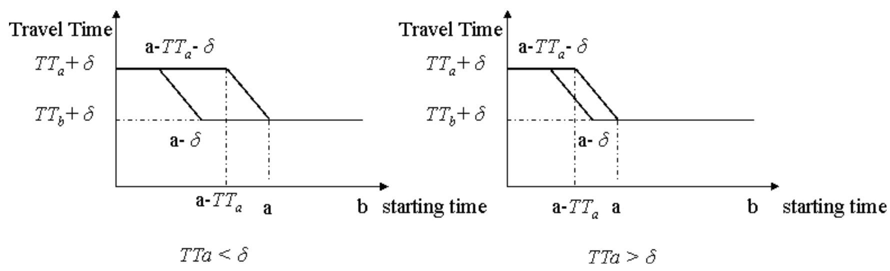

Figure 3.4 Illustration of the absorption effect AE

We analyze the absorption effect as a function of the starting times, which is relevant to the P-TDVRP. This is depicted by Figure **??** for both cases, $TT_a < \delta$ and $TT_a > \delta$, where $l_1 = TT_a - TT_b$ and $l_2 = \frac{TT_a - TT_b}{TT_a}\delta$. The absorption takes place within the interval $[a - TT_a - \delta, a]$. We distinguish between three cases depending upon the relative value of TT_a versus δ:

1. For the case where $TT_a < \delta$, the absorption is increasing for starting times within $[a - TT_a - \delta, a - \delta]$. In Figure **??**, this can be seen as the difference between the upper base of the parallelogram, the perceived flat line, and line of the actual time in the transient zone. For starting times within $[a - \delta, a - TT_a]$ the absorption is constant and is equivalent to $TT_a - TT_b$. For starting times within $[a - TT_a, a]$ the perceived travel time will follow the slope while the actual one will be TT_b, and thus the absorption will be decreasing.

2. For the case where $TT_a > \delta$, for starting times within $[a - TT_a - \delta, a - TT_a]$ the absorption is increasing, again this can be observed in Figure **??**, as the

difference between the flat line and the slope, however since $TT_a > \delta$ point $a - TT_a$ will occur before $a - \delta$. Starting times within the interval $[a - TT_a, a - \delta]$, will have a constant absorption since both lines, perceived and actual, will be in the transient zone with a constant difference equaling $\frac{TT_a - TT_b}{TT_a}\delta$. For the starting times within the interval $[a - \delta, a]$ the absorption will decrease. Dividing the area underneath both trapezoids with the planning horizon is equivalent to the value obtained in Equation (**??**).

3. For the case where $TT_a = \delta$ the shape in Figure **??** will become an isosceles triangle for both cases.

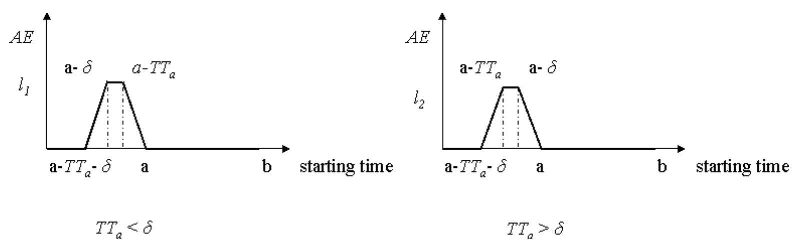

Figure 3.5 Absorption effect for one link

We note that a speed decrease results in a negative AE, i. e., adding a perturbation of size δ will, on average, result in a travel time that is greater than $TT + \delta$. We conclude that while a speed increase diminishes the effect of the perturbation, a speed decrease amplifies it. Hence, solutions to the perturbed problem should avoid situations that exhibit amplification of the perturbations.

3.5.2 The absorption effect with m links

In practice, more than one link will be traversed. Consequently, we are interested in determining the absorption effect, given a sequence of m links (and their corresponding travel times), assuming that each link is perturbed once. We will analyze the propagation of the absorption effect over these m links. More specifically, after perturbing the first link, to what extent the consecutive links are affected. We distinguish between two cases: one where the travel time profiles are identical for the different links and one where the travel time profiles are different per link. Clearly, the latter case is more realistic, but the first case is seen as a benchmark situation. In the remainder of this section, both cases will be discussed in detail.

Identical Travel Time profiles

Let us first focus on the simplified case where we have *two identical links* after each other. We have a route comprised of two links each following the same profile as the one described above in Figure **??**. The left side of Figure **??** illustrates the travel time for starting times on link 1, the upper left figure resembles Figure **??**. However the lower left part of Figure **??** depicts the travel time on link 2 given the starting times on link 1. Therefore the earliest start time on link 2 is after TT_a time units. When link 1 starts at $a - 2TT_a$ it arrives at link 2 at point $a - TT_a$, where the transient zone for link 2 starts. Link 2 remains in the transient zone till link 1 starts at $a - TT_a$ and then link 2 starts at point a with travel time TT_b. However, between starting times within $[a - TT_a, a]$, link 1 will be in the transient zone, while link 2 takes TT_b time units.

The right side of Figure **??** depicts the total travel time for links 1 and 2, as a function of the starting time on link 1. When the starting time is between zero and $a - 2TT_a$, the travel time on both links is TT_a. For starting times $[a - 2TT_a, a - TT_a]$ the travel time on link 1 is TT_a while the travel time on link 2 follows the slope pattern. For the interval $[a - TT_a, a]$, link 1 will follow the slope while link 2 will take TT_b. As of starting point a and onwards the travel time on both links is TT_b.

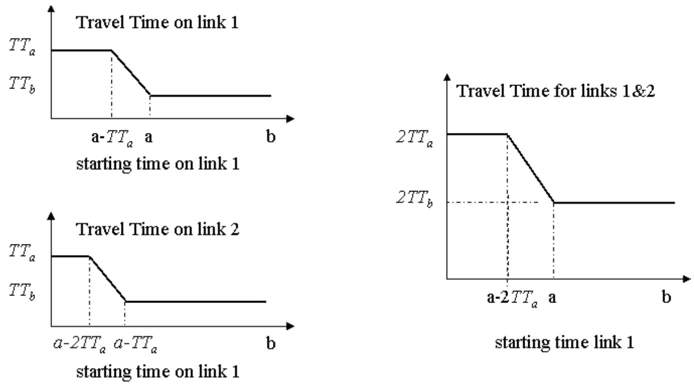

Figure 3.6 Travel time profile for two consecutive identical links as a function of the starting time of the first one

We define the average absorbtion affect for perturbing link i out of m as $\overline{AE}_{(i,m)}$, and the average absorption over all m links as $\overline{AE}_m$. The calculations for the average absorption effect in the case of perturbing link 1 out of 2 ($\overline{AE}_{(1,2)}$), are similar to those presented for the single link case:

$$\overline{AE}_{(1,2)} = 2\frac{\delta\left(TT_a - TT_b\right)}{b} \tag{3.8}$$

If link 2 is perturbed, link 1 is not effected and therefore the result from Equation (**??**) holds, thus perturbing each of the two links, as discussed in Section **??** will result in the following average absorption effect:

$$\begin{aligned}\overline{AE}_2 &= \frac{1}{2}\left(2\frac{\delta\left(TT_a - TT_b\right)}{b} + \frac{\delta\left(TT_a - TT_b\right)}{b}\right) \\ &= \frac{3}{2}\frac{\delta\left(TT_a - TT_b\right)}{b}\end{aligned} \tag{3.9}$$

Next we show the absorption effect as a function of the starting times for link 1 where $TT_a > \delta$, this is depicted by Figure **??**, where $l = \frac{TT_a - TT_b}{TT_a}\delta$. The region of starting times for which the absorption exists has increased with TT_a when compared to Figure **??**, and is between $[a - 2TT_a - \delta, a]$. A similar pattern could be generated for $TT_a < \delta$ with the same difference mentioned earlier for Figure **??**.

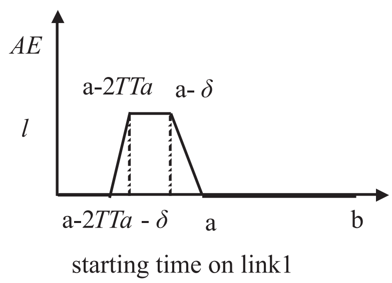

Figure 3.7 Absorption effect for two identical links

In the same manner, it is straightforward to show that, when the first out of m identical link is perturbed, that the absorption occurs in the region $[a - mTT_a - \delta, a]$. And that the following holds:

$$\overline{AE}_{(1,m)} = m\frac{\delta\left(TT_a - TT_b\right)}{b} \tag{3.10}$$

Similarly, the average absorption effect over m identical links ($\overline{AE}_m$), where each link is perturbed once, can be depicted by the following (where i represents the i^{th} link in the sequence):

$$\begin{aligned}\overline{AE}_m &= \frac{\delta \sum_{i=1}^{m}(m-i+1)(TT_a - TT_b)}{b \times m} \\ &= \frac{\delta(m+1)}{2b}(TT_a - TT_b) \end{aligned} \tag{3.11}$$

Different Travel Time profiles

We again look first into the case of a route comprised of two consecutive links with two different distances, following the same speed profile, i. e., each link has a different travel time profile. We analyze the propagating absorption effect: when the link 1 is perturbed to which extent are the consecutive links effected. We consider the case where link 1 follows the same profile as the one described Figure **??**, link 2 follows the same pattern with TT_c and TT_d instead of TT_a and TT_b respectively, however it has the same breaking point a. This follows from the assumption that a unique speed profile applies for both links. Assuming a unique speed profile is motivated by the fact that most roads exhibit a similar congestion pattern, (e. g., morning and afternoon rush hours occur at similar moments on most motor ways, **?**).

Figure **??** exhibits the total travel time of traversing link 1 and then later on link 2, we note that the arrival time on link 2 is starting at TT_a. For link 1 where starting time is between zero and $a - TT_a - TT_c$, the total travel time for both links is $TT_a + TT_c$. For starting times in $[a - TT_a - TT_c, a - TT_a]$ the travel time on link 1 is TT_a while link 2 is in its transient zone. For starting times $[a - TT_a, a]$ link 1 is in its transient zone while the travel time on link 2 is TT_d. As of starting point a and onwards the travel time for both links is $TT_b + TT_d$.

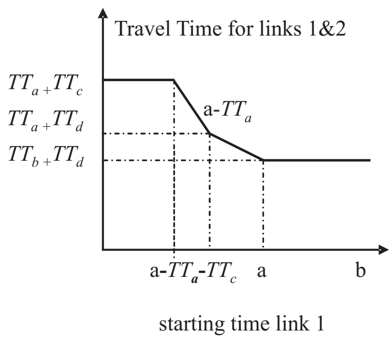

Figure 3.8 Travel time profile for two consecutive different links

The travel time augmented with the perturbation is then the following:

$$\begin{aligned}
T_{ijt} + \delta &= \frac{1}{b}\int_{o}^{a-TT_a-TT_c} (TT_a + TT_c + \delta)\, dt \\
&+ \frac{1}{b}\int_{a-TT_a-TT_c}^{a-TT_a} \left(\frac{TT_d - TT_c}{TT_c} t + TT_a + TT_d \right. \\
&\left. + (a - TT_a)\frac{TT_c - TT_d}{TT_c} + \delta\right) dt \\
&+ \frac{1}{b}\int_{a-TT_a}^{a} \left(\frac{TT_b - TT_a}{TT_a} t + TT_b + TT_d + a\frac{TT_a - TT_b}{TT_a} + \delta\right) dt \\
&+ \frac{1}{b}\int_{a}^{b} (TT_b + TT_d + \delta)\, dt \\
&= a(TT_c + TT_a - TT_d - TT_b) + \frac{TT_d TT_c - TT_a^2 - TT_c^2 + TT_b TT_a}{2} \\
&- TT_a TT_c + TT_d TT_a + b(TT_b + TT_d) + \delta \qquad (3.12)
\end{aligned}$$

However, when a perturbation is inserted the travel time would start to drop at point $a - \delta - TT_a - TT_c$, where the travel time follows the slope of link 2, and consecutively starting at point $a - \delta - TT_a$ travel time follows the slope of link 1. Thus the actual perturbed travel time is:

$$\begin{aligned}
TT_{ijt,\delta} &= \frac{1}{b}\int_{o}^{a-TT_a-TT_c-\delta} (TT_a + TT_c + \delta)\, dt \\
&+ \frac{1}{b}\int_{a-TT_a-TT_c-\delta}^{a-TT_a-\delta} \left(\frac{TT_d - TT_c}{TT_c} t + TT_a + TT_d \right. \\
&\left. + (a - TT_a - \delta)\frac{TT_c - TT_d}{TT_c} + \delta\right) dt \\
&+ \frac{1}{b}\int_{a-TT_a-\delta}^{a-\delta} \left(\frac{TT_b - TT_a}{TT_a} t + TT_b + TT_d + (a-\delta)\frac{TT_a - TT_b}{TT_a} + \delta\right) dt \\
&+ \frac{1}{b}\int_{a-\delta}^{b} (TT_b + TT_d + \delta)\, dt \\
&= \delta(TT_b - TT_c - TT_a + b + TT_d) + \frac{TT_d TT_c - TT_a^2 - TT_c^2 + TT_b TT_a}{2} \\
&a(TT_a + TT_c - TT_d - TT_b) - TT_a TT_c \\
&+ b(TT_b + TT_d) + TT_d TT_a \qquad (3.13)
\end{aligned}$$

The average absorption effect, in the case of perturbing link 1, can be obtained by subtracting Equation (??) from Equation (??):

$$\overline{AE}_{(1,2)} = \frac{\delta\left(TT_a + TT_c - (TT_b + TT_d)\right)}{b} \tag{3.14}$$

The absorption affect occurs in within the interval $[a - TT_a - TT_c - \delta, a]$. The shapes of the trapezoids, that depict the AE, depend on the relation between TT_a and δ, as well as on the relation between TT_c and δ. Similarly, for a sequence of links $\{1, \ldots, i, \ldots, m\}$ where each link follows the same speed profile as the one described in Figure **??**, i.e, with the same breaking point a, while each link has a high travel time TT_{a_i} and low travel time TT_{b_i}. The resulting average absorption effect, where each link is perturbed once, over all n is given by Equation (??). This is inferred by generalizing Equation (??) to a sequence of links.

$$\overline{AE}_m = \frac{\delta \sum_{i=1}^{m} i\left(TT_{a_i} - TT_{b_i}\right)}{bm} \tag{3.15}$$

We note that the above analysis is relevant for the case where $TT_a < a$. The travel time functions used in the Section **??** were calculated as integrals over speeds, and the figures presented earlier adhere to this. Equivalent figures would differ for the case where $TT_a > a$. However, in principle the same absorption behavior will occur: a positive AE with a decrease in travel time and a negative AE with an increase in travel time.

3.6. Experimental settings

In Section **??**, we introduce our speed profile. In Section **??**, we discuss the Tabu search solution procedure. In Section **??**, the results are presented and discussed.

3.6.1 The speed profile used

We used a speed profile such that, during the day, two periods with a relatively low speed exist, while throughout the rest of the day the speed is relatively high (Figure **??**). Many roads incur such a morning and an afternoon congestion period (see, e. g., ?). We use a speed profile, derived from data collected for a Belgian highway, as shown in Figure **??** where the right side of the figure depicts the travel time function for each start time for an arbitrary link, under the following conditions: $v_2 = v_4$ and $v_5 > v_3 > v_1$ (since this complied with the data on which the profile was constructed). We observed the following average speeds $v_1 = 95$ km/hr for the period 0:00-6:00,

$v_2 = 75$ km/hr for the period 6:00-9:00, $v_3 = 100$ km/hr for the period 9:00-16:00, $v_4 = 75$ km/hr for the period 16:00-19:00, $v_5 = 105$ km/hr for the period 19:00-0:00.

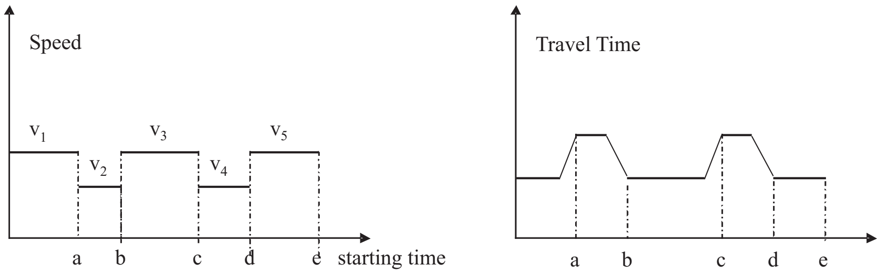

Figure 3.9 Travel time profile for two congested profile

Both congestion regions $[a, b]$ and $[c, d]$ start with a negative AE and end with a positive one, assuming a uniform distribution on the start times $[0, e]$, the average absorption should be $\frac{\delta\left(TT_{v_1} - TT_{v_2} + TT_{v_2} - TT_{v_3}\right)}{e} + \frac{\delta\left(TT_{v_3} - TT_{v_4} + TT_{v_4} - TT_{v_5}\right)}{e}$, which is positive with the chosen speed profile. However, the absorption affect for a discrete solution would depend on the sequence of links in question.

For simplicity, we assumed that the speed profile described above holds for all links in the graph. We assumed an overtime penalty of $p = 9$ and a shift length $w = 600$ minutes. We also impose the constraint that each tour must end before midnight. The solutions for both TDVRP and P-TDVRP are obtained via a tabu search procedure that optimizes on a number of starting times.

3.6.2 The Tabu Search procedure used

A solution strategy based on local search is proposed. According to **?**, local search is a solution process that tries to improve a given initial solution by making relatively small changes in several steps in the solution space. The quality of the solutions is then determined with the target function of the problem. In this chapter, tabu search is used for obtaining solutions. Tabu search was first introduced by **??**. It makes use of adaptive memory to escape local optima. The method has been extensively used for solving VRPs (see, for example, **??**, and **?**).

In our implementation, which is similar to **?**, time-dependent travel time is added as a feature. For the P-TDVRP, we incorporate the perturbed target function. The starting time of traversing a route is an important variable for a time-dependent environment, since this moment dictates the total travel time of the route. In order to take starting times into account each solution is evaluated for four hourly starting times between 6:00 and 9:00. Our implementation of the P-TDVRP setting involves additional computation time since each solution is now evaluated m times, where m is the number of links that comprise a solution. To explore the added value of having an additional vehicle when optimizing the perturbed setting, we enable the use of one extra vehicle, in comparison to the TDVRP. No extra costs are associated with this additional vehicle. This sheds light on the added value of having an additional vehicle when optimizing the perturbed setting.

The initial solution is constructed by the nearest neighbor heuristic. A complete neighborhood of a solution is evaluated by a combination of 2-opt and swap moves for the 15 closest neighbors of a node. The tabu list was randomly chosen within $[22, 34]$. We use an aspiration criterion similar to **?**. This criterion revokes the tabu status of a move if it yields a solution with lower costs.

Similar to **?**, the target function includes penalty components for infeasibility constraints, one for exceeding the truck capacity and another for exceeding the planning horizon ,i.e, midnight.

3.7. Propositions and Results

In order to validate the existence of low-cost solutions under the perturbed setting and to explore their structural properties, we experimented with sets from **?**. These sets have customers ranging between 31 and 79. To achieve more realistic travel times, the coordinates of these sets were multiplied by 4.5. To explore the existence of the behavior discussed in section **??**, we formulate a set of propositions in the following section. The computational results will be presented in Section **??**.

3.7.1 Propositions

We have clustered our propositions into three main groups in a hierarchial manner. The first group deals with the overall performance of solutions S^* and Z^*. Remember that S^* is the best solution found under the TDVRP optimization, while Z^* is the best solution found with the P-TDVRP optimization. The second group handles the solutions by the different structure of the routes that make up the solutions S^* and Z^*. The third group focuses on the individual links that comprise the routes.

Aggregate propositions

- P_{A1}: $G - C > 0$.
  This proposition states that the trade-off will be positive. In essence, the trade-off was defined as the difference between gains and costs, given by:

$$G - C = [F_{pert}(S^*) - F_{pert}(Z^*)] - [F_{norm}(Z^*) - F_{norm}(S^*)] \tag{3.16}$$

  We predict that not only there exist solutions that will outperform the normal solution in the perturbed setting, but that the gain resulting from these solutions outweighs the cost of using them.

- P_{A2}: $AE(Z^*) > AE(S^*)$.
  The absorption effect for the solutions resulting from the perturbed setting (Z^*) is higher than that of the normal solutions (S^*) when perturbed. The absorption effect as defined in Section **??** is the difference between the perceived travel time and the actual travel time when perturbed. The above proposition implies that the perturbed solutions will result in higher absorptions than the normal ones. The underlying assumption behind this is that the performance of Z^*, when subjected to disruption, will outperform S^* due to more absorption abilities. This is due to the fact that F_{pert} includes perturbations that impact the absorbtion.

Route based propositions

In both the normal and the perturbed settings, routes are penalized by p for every minute exceeding the shift length w. During the optimization of P-TDVRP, the routes are subjected to a certain perturbation which in return will increase the travel time, albeit not perse by the perturbation size due to the absorption effect (see Section **??**). As a consequence, we expect to see structural changes in the individual routes resulting from the perturbed setting.

More specifically, we expect routes in the perturbed setting to be more sensitive to the over-time penalty, i.e., being more conservative with respect to the w threshold. Consequently, it is expected that the long tours will become shorter to protect themselves better against the possible delays. Shorter routes will become longer or more tours will be added to compensate for this effect. This is expected to result in a leveling of the tour lengths. In order to verify this, we construct the following propositions.

- P_{R1}: The average route length will increase in Z^* compared to the one in S^*. This is motivated by the following relation $F_{norm}(S^*) \leq F_{norm}(Z^*)$, since the F_{norm} relies heavily on the total travel time, without perturbations. The fact that the perturbed solutions will be equivalent or outperformed by the normal ones implies an increase in the total travel time.

- P_{R2}: The maximum route length will be lower for Z^* compared to its equivalent in S^*.
  We expect the maximum route length to decrease in the P-TDVRP. Since the longest tour is the closest to the shift length boundary, subjecting it to perturbations may lead to crossing this threshold . In order to reduce the impact of the perturbations, the longest tour will decrease.

Link based propositions

In Section **??**, we discussed situations where a positive or a negative absorption may occur. The former implies that the perturbation is translated to additional travel time that is less than the perturbation size. Similarly, a negative absorption depicts situations where the additional travel time, resulting from perturbation, is higher than the perturbation itself. We have shown that while a positive absorption occurs in zones where the speed increases, the negative one occurs in zones where speed decreases.

In Section **??**, the average absorption effect was expressed on the basis of assuming a uniform distribution on the starting times. However, a solution to the TDVRP is comprised of a sequence of links, each traversed in a single discrete point, that is determined either by its predecessor link or by the starting time of the route (for the first link on the route). Thus, while on average the absorption effect can be calculated over a distribution on the starting times, the actual absorption, whether positive or negative, will depend on the discrete solution instance.

We note that for the speed profile depicted in Figure **??**, it is highly likely that the routes will start from point b onwards. This is due to the fact that both the morning and afternoon congestion periods are of the same length, i. e. three hours, while $v_1 < v_3 < v_5$. Thus for a route ending before midnight it is certainly better to commence after b. Our experiments have shown that this holds for most cases. Thus in the subsequent propositions we exclude zones prior to b since they are irrelevant.

In order to gain insights with regards to the dispersion of the allocation of links in the different speed zones, we partitioned the planning horizon into nine different speed zones (e. g., in Figure **??** “v_{45}” means on the slope between v_4 and v_5). The following propositions relate the theoretical findings with the chosen experimental setting. P_{l1}

and P_{l2} predict the spread of solutions with respect to the number of links traversed in specific speed zones. We always compare the P-TDVRP solutions to the TDVRP ones.

- P_{l1}: More links are traversed in speed zones v_4 and v_{45}.
  This stems from the fact that a positive absorption occurs where there is a speed increase. In the experimental setting this is illustrated in speed zones v_4 and v_{45}. In TDVRP the absorption effect does not play a role. Thus we predict that the number of links traversed in the speed zones in question will be higher in the P-TDVRP setting.
- P_{l2}: Less links are traversed in speed zones v_3 and v_{34}.
  Similar to P_{l1}, negative absorption occurs with a speed decrease, which happens in speed zones v_3 and v_{34}. Again this would be irrelevant to the normal optimization. Thus, the number of links in these speed zones would be lower in the P-TDVRP.

3.7.2 Results

We experimented with perturbations sizes $\delta \in \{20, 50, 100, 200\}$, the values are in minutes. The choice was based on discussions with carrier companies. The runs were conducted on Pentium 4 PC with 3.00 GHz and 2 GB of RAM. We partition the results section similar to the structure of the propositions: from aggregate results over route based results and finally to link based results.

Aggregate results

Table **??** gives the trade-off for the different sets analyzed in columns (2-5). It also gives the runtime required to attain each of the solutions in columns (6-10). A positive value of the trade-off is in favor of the P-TDVRP solution found, while a negative one shows that the solution obtained from the TDVRP is better. A trade-off of zero indicates that the solution found under the TDVRP optimization is identical to the one found under the P-TDVRP. Lines (2-4) from the bottom show the number of instances, out of 27, where the trade-off was greater, smaller or equal to zero respectively.

From Table **??**, we first observe that 69 out of the 108 experimental instances have a trade-off value that does not equal zero, this indicates that in nearly two thirds of the cases there exists a solution that outperforms the normal one in the P-TDVRP setting. The number of instances with a trade-off different from zero rises with the perturbation size. Furthermore, Table ?? exhibits the run times for all experimental sets in minutes. We observe that the run times are predominantly influenced by the instance size, and are by far greater than the run times for the normal optimization.

	Trade-off ($G-C$)				Calculation times (in minutes)					Category
Set	$\delta = 20$	$\delta = 50$	$\delta = 100$	$\delta = 200$	S^*	$\delta = 20$	$\delta = 50$	$\delta = 100$	$\delta = 200$	$\{A, B, C\}$
32 k5	0.0	727.4	9646.3	17164.6	0.7	23.7	50.8	43.6	12.2	A
33 k5	0.0	0.4	33.3	100.8	0.7	16.8	31.4	58.6	25.6	A
33 k6	0.0	0.2	1.4	-153.4	0.9	32.3	59.1	79.4	25.5	C
34 k5	0.0	0.0	-10.5	-6.0	0.9	23.1	45.1	68.1	25.4	B
36 k5	0.0	966.6	3969.0	10246.4	1.0	33.7	61.9	61.3	20.9	A
37 k5	0.0	0.0	0.0	62.9	1.0	30.0	83.5	25.7	19.4	A
37 k6	1.1	0.0	25.3	7346.5	1.5	33.2	65.7	66.1	55.5	A
38 k5	0.0	0.0	44.3	217.7	1.2	67.9	73.8	47.0	51.4	A
39 k5	0.6	570.5	5288.7	21511.0	1.6	16.4	82.8	27.4	28.2	A
39 k6	2.0	1.0	-25.4	-22.6	0.7	15.2	25.5	50.9	26.1	C
44 k6	0.0	-4.7	0.0	-191.9	1.2	105.0	182.8	193.0	71.7	B
45 k6	1.7	4.4	40.4	113.9	1.6	55.1	109.0	88.3	93.0	A
45 k7	0.0	-28.0	6231.0	15195.6	1.8	135.9	204.1	163.4	179.6	C
46 k7	0.0	0.0	-9.1	-1.3	2.2	91.5	157.1	152.2	23.5	B
48 k7	0.0	23.8	67.6	115.6	1.4	52.7	113.1	139.9	49.2	A
53 k7	-6.3	0.0	16.3	24.3	3.0	114.8	173.7	280.8	219.6	C
54 k7	0.0	26.0	19.7	100.5	7.1	236.0	135.0	414.3	189.8	A
55 k9	0.0	0.0	0.0	0.0	6.0	323.7	246.5	341.7	228.2	A
60 k9	0.0	0.0	0.0	0.0	12.8	552.8	479.4	342.0	304.7	A
61 k9	0.0	0.0	0.0	0.0	8.1	367.5	238.4	496.8	464.6	A
62 k8	12.3	2350.6	9147.0	16977.1	6.9	487.9	193.5	676.4	210.5	A
63 k9	0.0	1753.7	1633.2	3202.1	9.9	895.4	592.3	736.3	674.9	A
63 k10	0.3	1.4	0.0	137.7	16.4	841.2	590.9	707.0	275.8	A
64 k9	300.2	1649.6	0.0	14778.1	18.6	838.0	329.6	510.0	340.9	A
65 k9	0.7	0.0	0.0	8.6	11.2	615.3	697.8	551.0	281.3	A
69 k9	0.1	0.4	3.7	19.7	13.3	1135.6	973.0	473.0	840.4	A
80 k10	0.0	0.0	1355.1	5502.4	30.8	1744.5	2680.5	833.8	1020.4	A
$\# > 0$	9	14	16	19					# of A = 20	
$\# = 0$	17	11	8	3					# of B = 3	
$\# < 0$	1	2	3	5					# of C = 4	

Table 3.1 Trade-off and Calculation times for perturbation sizes $\delta \in \{20, 50, 100, 200\}$

Since there are different trends in terms of trade-off, we classified the sets into the following three categories:

$$A \; : \; \text{instances where} \quad G - C \geq 0 \quad \forall\, \delta$$
$$B \; : \; \text{instances where} \quad G - C \leq 0 \quad \forall\, \delta$$
$$C \; : \; \text{instances where} \quad \exists\, \delta \text{ with } G - C > 0 \text{ and } \exists\, \delta \text{ with } G - C < 0$$

This classification can be found in the last column of Table **??**. Category A is defined as one where the solutions from P-TDVRP have a trade-off that is greater or equal to zero for all four perturbations sizes used. This means that the perturbed solution is either better or equivalent to the normal one in all cases. Generally, category A exhibits a trend of the trade-off rising with the perturbation size, (e. g. sets 32k5, 36k5, and 62k8). In these examples, one might classify the sets as hyper-sensitive to the disruption size. Contrary to Category A, Category B is one where the P-TDVRP solutions are either equivalent or worse than the normal ones. We note that the negative values associated with this category are relatively small in absolute terms. If there is a tie between A and B we choose A. Finally Category C, is one where the sets are inconsistent for the four perturbation sizes, this means that for some perturbation sizes the trade-off is positive while for others it is negative. Category A dominates the experimental sets with 20 observation out of 27, this implies that in many cases we are better off using Z^* instead of S^*.

Proposition P_{A1} states that the average trade-off is greater than zero. We observe that out of the 108 experiments only 11 hade a negative trade-off, while 58 had a positive trade-off. Thus, we conclude that P_{A1} holds for the entire experimental set.

Observing the results for each of the perturbation size separately. For perturbation size 20 Table **??** shows that 17 out of the 27 resulted in a trade-off of zero. However, the number of instances with positive trade-off increases with the perturbation size. This implies that for relatively small perturbations the results are not substantially different than zero while for larger ones the differences are more significant. We note that some of the gains are high since evaluating $F_{pert}(S^*)$ includes the penalties of infeasibility constraints, thus if the perturbed solution exceeds the planning horizon it is penalized.

Next, we look into the attained absorption effect values (AE) in Table **??**. As in Section **??** the AE is calculated as the difference between the perceived perturbed travel time and the actual travel time. In Table **??** columns (2-5) exhibit the AE for each of the solutions resulting from the P-TDVRP, while columns (6-9) exhibit the AE for subjecting the normal solutions from the TDVRP to four different perturbation sizes. Thus the right side of the table is dedicated to the AE as attained by the perturbed optimization while the right side of the table shows the performance of the normal solutions, one for each set, in terms of AE for the perturbation sizes.

Out of the 108 instances in the P-TDVRP optimization setting there were 96 positive ones. The corresponding number for the TDVRP is 79. When looking at the average AE over the entire 108 instances, the solutions from the P-TDVRP setting have an average AE of 30.7 min while the solutions from the TDVRP setting have an average of -139.1 min, this implies that in many cases the TDVRP solutions resulted in extremely negative AE values. This can be explained by the fact that the optimization procedure for TDVRP does not consider perturbations. In total, 107 observations in the P-TDVRP setting had a higher or equal absorption effect than those of the TDVRP ones.

	AE for $Z*$				AE for S^*			
Set	$\delta = 20$	$\delta = 20$	$\delta = 100$	$\delta = 200$	$\delta = 20$	$\delta = 50$	$\delta = 100$	$\delta = 200$
32 k5	11.9	34.8	70.9	93.3	11.9	-46.3	-1024.4	-1839.4
33 k5	17.5	43.7	87.1	147.5	17.5	43.7	79.8	127.5
33 k6	17.4	43.4	86.0	135.6	17.4	43.4	85.8	131.2
34 k5	17.4	43.2	86.5	139.1	17.4	43.2	82.7	126.7
36 k5	8.3	22.8	36.4	65.0	8.3	-87.6	-414.6	-1092.1
37 k5	17.5	41.8	71.5	150.5	17.5	41.8	71.5	116.6
37 k6	16.1	33.4	58.3	89.3	15.3	33.4	56.9	-734.0
38 k5	17.4	43.6	80.9	147.7	17.4	43.6	74.8	118.3
39 k5	12.1	28.5	45.0	81.1	12.0	-40.8	-548.0	-2331.3
39 k6	15.0	33.2	64.5	105.7	13.8	29.1	52.5	89.5
44 k6	10.0	30.0	51.2	114.9	10.0	27.4	51.2	92.5
45 k6	14.8	35.7	71.9	119.6	14.6	35.2	66.5	105.6
45 k7	12.0	27.5	50.6	75.0	12.0	19.7	-648.7	-1641.1
46 k7	17.3	42.0	80.7	136.8	17.3	42.0	79.2	129.3
48 k7	10.0	33.5	56.0	83.1	10.0	22.1	41.0	65.4
53 k7	16.7	32.8	72.4	124.2	11.5	32.8	66.6	111.3
54 k7	12.8	38.0	62.0	126.8	12.8	26.6	46.9	79.1
55 k9	17.6	44.0	85.9	157.2	17.6	44.0	85.9	157.2
60 k9	-43.7	-20.4	10.4	56.7	-43.7	-20.4	10.4	56.7
61 k9	17.6	44.0	88.0	159.6	17.6	44.0	88.0	159.6
62 k8	3.6	4.6	21.8	17.9	1.7	-265.7	-1003.7	-1926.2
63 k9	-132.1	-124.5	-113.1	-9.8	-132.1	-324.4	-312.3	-552.8
63 k10	17.1	39.6	68.9	121.5	17.0	39.6	68.9	116.0
64 k9	-9.1	5.8	-854.3	39.7	-43.2	-182.3	-854.3	-1594.0
65 k9	16.4	39.0	76.7	142.3	16.3	39.0	76.7	143.8
69 k9	17.6	43.8	85.4	146.5	17.5	43.8	84.9	144.3
80 k10	-255.8	-316.8	-148.2	-163.7	-255.8	-316.8	-304.3	-804.1
Average	-4.0	13.6	16.8	96.4	-5.6	-21.9	-142.2	-386.8

Table 3.2 Absorption effect for perturbation sizes $\delta \in \{20, 50, 100, 200\}$ for both optimization settings

When looking at the different perturbation sizes, we note that, while in the P-TDVRP the AE is increasing with the perturbation size, in the TDVRP it is decreasing. Furthermore, in the TDVRP the AE is on average always negative, whereas in the P-TDVRP it is only negative for perturbation size 20. Proposition P_{A2} suggests that the AE for Z^* is higher than for S^*. We observe perturbations size 50, 100 and 200 the AE is considerably larger for Z^* compared to S^*. However, this is not the case for perturbation size 20. This can be explained by the fact that in this case 17 out of the 27 observations are identical, as seen by Table **??**.

Route based results

In order to validate the relevant propositions, we construct Table **??**. The data is based on the travel time as in the normal setting, i. e. without perturbations. This is motivated by learning more about the structure of the routes that make up solutions in the perturbed optimization. As a consequence, we compare five different solution sets: one for the normal optimization and four for the perturbation sizes.

	Average route length					Maximum route length				
Set	S^*	$\delta = 20$	$\delta = 50$	$\delta = 100$	$\delta = 200$	S^*	$\delta = 20$	$\delta = 50$	$\delta = 100$	$\delta = 200$
32 k5	546.4	546.4	546.7	**473.0**	**471.9**	597.0	597.0	594.7	597.6	603.3
33 k5	367.1	367.1	367.1	370.3	378.4	526.4	526.4	526.4	469.8	489.8
33 k6	341.1	341.1	341.1	341.2	373.2	482.9	482.9	482.9	483.2	448.1
34 k5	368.9	368.9	368.9	372.5	384.6	503.9	503.9	503.9	460.6	480.6
36 k5	529.8	529.8	530.0	**428.3**	**439.6**	605.8	605.8	608.6	633.0	634.5
37 k5	320.1	320.1	320.1	320.1	*409.1*	557.2	557.2	557.2	557.2	527.3
37 k6	449.3	450.0	449.3	447.1	**451.0**	600.5	597.6	600.5	602.4	606.0
38 k5	410.9	410.9	410.9	**343.3**	**349.3**	540.8	540.8	540.8	523.6	548.1
39 k5	473.6	473.6	476.3	475.9	413.4	602.3	602.3	605.1	605.1	610.6
39 k6	393.5	394.2	396.5	404.5	410.3	592.7	592.7	589.9	585.2	582.5
44 k6	404.2	404.2	407.4	404.2	442.0	601.9	601.9	599.0	601.9	587.9
45 k6	384.8	384.8	384.8	385.4	385.9	595.5	595.5	595.5	595.3	592.5
45 k7	505.3	505.3	512.9	510.4	**470.9**	600.8	600.8	592.1	597.8	602.6
46 k7	381.5	381.5	381.5	383.1	391.5	556.7	556.7	556.7	555.2	559.7
48 k7	454.9	454.9	460.5	459.7	456.4	598.6	598.6	587.1	590.0	597.8
53 k7	403.3	408.8	403.3	407.6	415.6	602.7	586.7	602.7	596.8	593.9
54 k7	502.0	502.0	507.5	**446.6**	499.4	599.4	599.4	586.6	590.9	598.1
55 k9	326.1	326.1	326.1	326.1	326.1	494.8	494.8	494.8	494.8	494.8
60 k9	430.7	430.7	430.7	430.7	430.7	659.3	659.3	659.3	659.3	659.3
61 k9	298.2	298.2	298.2	298.2	298.2	470.0	470.0	470.0	470.0	470.0
62 k8	517.9	518.2	518.9	518.9	517.8	607.0	606.8	609.9	609.9	612.0
63 k9	525.5	525.5	526.7	531.6	**475.9**	745.8	745.8	748.7	751.5	745.8
63 k10	383.5	383.5	383.5	383.5	380.1	586.0	586.0	586.0	554.7	556.9
64 k9	504.2	504.5	506.0	504.2	**426.7**	623.9	624.2	625.0	623.9	656.7
65 k9	331.5	331.5	331.5	331.5	352.8	591.0	591.0	591.0	591.0	512.1
69 k9	360.7	360.7	360.7	360.7	359.5	495.2	495.2	495.2	495.2	454.1
80 k10	545.0	545.0	545.0	540.6	528.9	752.2	752.2	752.2	757.9	796.5
Average	424.5	424.7	425.6	414.8	416.3	584.8	584.1	583.8	579.8	578.6

Table 3.3 Average and maximum route length for perturbation sizes $\delta \in \{20, 50, 100, 200\}$ for both optimization settings

The left side of of Table **??** depicts the average route length for each instance, while the right side of the table shows the maximum observed route length (all in travel times). Bold numbers indicate that an extra vehicle was used in P-TDVRP. Italic numbers indicate that one vehicle less was used in P-TDVRP.

P_{R1} states that the average route length will increase in Z^* compared to the one in S^*. In the case of perturbation size 20, the average route length is higher than that of the normal solutions. P_{R1} is more evident for perturbation size 50, the average of Z^* is higher that S^*.

The decrease in the average tour length for the higher perturbation settings (100 and 200) is due to an increase in the number of vehicles used. The instances are indicated in bold in the left side of Table **??**. No extra vehicle was used in any of the 27 solutions for perturbation size 20 or 50. However for perturbation size 100, there was a need for an extra vehicle in 4 instances, while for perturbations size 200 there was a need for 7 extra vehicles. This implies that the number of vehicles needed to protect from perturbations rises with the perturbation size, and as a consequence the average route length might decrease. The table shows inconsistent behavior for five values, where the average route length is lower for Z^* than for S^* and the number of vehicles is the same for both solutions. The gap between these solutions is less than 1% for four out of the five instances, and is at most 3% for set 80k10 with perturbation 200. These inconsistencies can be avoided by running the algorithm again for TDVRP, while setting the initial solution to Z^*. To ensure further consistency, S^* may also be considered as an initial solution to P-TDVRP.

We note that for perturbation size 200 in set 37k5 there was a need for one vehicle less, indicated by italic in Table **??**, the reason for that is the existence of a very short route, in TDVRP, that was omitted in the P-TDVRP optimization. For perturbation sizes 100 and 200 the average route length is smaller than in S^*.

We have experimented with different values of the overtime penalty of p (in the presented experimental setting $p = 9$). We observed that lower values of this parameter do not substantially influence the Trade-off values. However, they may influence the use of additional vehicles. Thus, for lower values of p the P-TDVRP solution may not require additional vehicles.

The data for P_{R2} can be observed in the right part of the table. In general the average maximum route length decreases with the increase in the perturbation size. P_{R2} is observed for perturbation size 50 and 100. The maximum route length in Z^* is not higher than the one in S^* in all instances. This can be explained by noting that if the travel time is increasing, then the maximum length might also increase. Additionally the maximum route can become longer in the P-TDVRP and as a consequence other routes become shorter. Thus resulting in a situation where the long route is sacrificed for the benefit of other routes that are more protected against perturbations.

Link based results

In order to gain insight into the dispersion of the solutions with respect to the allocation of links in the different speed zones, we generate Table **??**. Remember that we partitioned the planning horizon into nine different speed zones. The basis for P_{l1-2} is that the distribution of the links will follow that of the logic presented in Section **??**; positive values of the absorption effect occur with speed increases while negative ones occur with speed drops.

The data required for observing the relevant propositions is summarized in Table **??** We present the data for the average number of links traversed in the different time zones over the 27 sets. We note that much of the links are traversed in speed zone v_3 since it is relatively large.

	Average number of links				
Set	S^*	$\delta = 20$	$\delta = 50$	$\delta = 100$	$\delta = 200$
Average number of links in $v_3 + v_{34}$	42.0	40.1	40.8	47.0	50.7
Average number of links in $v_4 + v_{45}$	3.9	10.7	12.6	11.8	7.1

Table 3.4 Average number of links and distance traversed by speed zone for perturbation sizes $\delta \in \{20, 50, 100, 200\}$

P_{l1} states that more links are traversed in speed zones v_4 and v_{45} in Z^* than in S^*. For the case of perturbation size 20 and 50 this is true, since the average number of links in these zones has increased in solutions of the P-TDVRP. For perturbation size 100 and 200, the average of number of links traversed in speed zones v_3 and v_{34} is significantly higher. The main reason for this is due to the large disruption size. When inserting a delay the links are compelled to start at a later point. If the delay is large the links that are likely to display the absorption are ones in speed zone v_3, which is earlier than the zones assumed by P_{l1}. This is further confirmed by the relatively high AE values shown in Table **??**, which indicate that this is the particular structure that produces better P-TDVRP solutions. Another reason is the existence of the planning horizon limitation, by which routes are not allowed to exceed midnight. When solutions are subjected to high perturbations routes are likely to exceed midnight and thus the routes in P-TDVRP will tend to end much sooner, in order to account for this. Yet another explanation is the fact that some additional vehicles are used (for perturbation size 100 and 200) which shorten the average route length.

Due to the reasons mentioned above, the number of links traversed in v_4 and v_{45}

for perturbation sizes 100 and 200 are significantly lower than the solutions resulting from the TDVRP. We observe that for perturbation size 20 and 50 P_{l2} exhibits the same behavior as P_{l1}.

3.8. Conclusions and further research

We addressed the problem of scheduling vehicles encountering delays at docking stations. We accounted for daily congestion by travel time functions that adhere to the FIFO assumption. A framework for obtaining solutions that better cope with unanticipated delays is presented. We showed theoretically how solutions could change in order to cope with these disruptions. We presented an experimental context for the problem, formulated propositions that support the framework. We then tested these hypotheses based on the experimental results.

In terms of methodology, this chapter applied a perturbation approach on a discrete space in a tabu search setting. The results showed that this approach was capable of finding solutions that were superior to normal ones under perturbations. While the presented approach was applied in a Tabu Search setting, its advantage is that it can be easily applied to a wide range of existing solution methods. Hence, carrier companies can adopt the approach without the need of drastic structural changes to their existing systems. The most important drawback of the approach is the extensive computational effort it requires. However, since it is designed for off-line scheduling, meaning that information with respect to customer demand is known, moreover rerouting is not possible, one could argue that calculation time is operationally affordable.

We showed that the solution structure changes to accommodate for the perturbations. In general these changes involved making use of the absorption effect, which reduces the resulting damage from unanticipated delays in the starting times of the links. As the TDVRP mainly optimizes travel times, it does not take the absorption into consideration. Solutions resulting from this optimization will therefore perform poorly under perturbations if their structure yields negative absorption. The experiments presented in this chapter support this notion. The results showed that the perturbation size is a highly influential parameter. This is shown by the trade-off, which increased with the perturbation size. From this result, we infer that the importance of the P-TDVRP increases with the perturbation size. Consequently planners should account for this. The results however became less straightforward when we analyzed the composition of the routes. In our setting we enabled the use of an extra vehicle. While this option was not needed for lower perturbation sizes, it was so for higher ones. This sunk expenditure is justified by the gains it brings. When there is no need for extra vehicles, the routes in the P-TDVRP are leveled, meaning that longer tours become shorter and as a compensation shorter ones become longer, this leads to additional travel time which is balanced by the tradeoff. Furthermore we have shown that the perturbation size effects the distribution of the links, higher perturbations push links, further away from the congestion period.

The P-TDVRP considered a single disruption per scenario. The proposed methodology can fit situations with multiple disruptions. These disruptions may be fitted in the proposed objective function. However, accounting for interactions between multiple disruptions may result in particularly long run times. Thus, further research on multiple setting may be conducted by sampling a limited number of scenarios.

Further research could include converting perturbations into road accidents, or other unexpected delays during travel time (as in Chapter 4). This would require adjustments with respect to the location of the perturbation and further analysis of the resulting AE. We experimented with constant perturbation sizes. Since disruptions are stochastic, for example low perturbations have a high probability while higher perturbations have a lower one, it is worthwhile investigating multiple disruption scenarios.

Chapter 4

Self-Imposed Time Windows in Vehicle Routing Problems

4.1. Introduction

Many small-package shipping companies provide their customers with a time window for delivery and display this in their online tracking system. UPS, for instance, shows information on the delivery time window of orders for DELL computers. Obviously, once a time window is quoted to the customer, the carrier company wants to service the client within this window and so this should be reflected in the carrier's routing decisions. The described environment is clearly distinct from both the classic Vehicle Routing Problem (VRP) as well as from the VRP with Time Windows (VRPTW). It is different from the VRP since the VRP objective is to minimize the operational costs (e. g., distances or travel times **?**). The VRPTW, on the other hand, does consider time windows but assumes they are exogenous, i. e., imposed by the customer **??**. As a consequence, the VRPTW imposes restrictions on the specific arrival times at each customer, while maintaining the objective of minimizing operational costs.

Our problem at hand considers time windows but treats them as endogenous to the VRP model. Specifically, the carrier company assigns customers to vehicles, sequences the customers allocated to each vehicle, and sets the time windows in which it plans to serve the customers. In the remainder of this chapter, we will refer to the described problem as the Vehicle Routing Problem with Self-Imposed Time Windows (VRP-SITW). The term 'self-imposed' refers to the fact that the carrier company selects the time windows by itself, independently of the customer. Once the time windows are quoted to the customer, however, the customer should be serviced within the window. As such, the VRP-SITW conceptually lies in between the VRP and the VRPTW. We assume that service cannot start before the time window, leading to waiting in case of early arrivals. Furthermore, late arrivals are permitted but penalized proportionally to their tardiness. Drivers have a fixed shift length and are paid a fixed amount per day. Finally, disruptions in traveling time may occur between each two customers. This mainly reflects accidents, weather condition, vehicle breakdown or road works. One natural way to protect schedules against this uncertainty is to include time buffers (see, for instance, **?** for a similar logic in a production environment). Inspired by the scheduling literature, we propose a buffer allocation model that inserts slack time into the schedule to cope with possible delays. Our solution framework relies on the tabu search heuristic for assigning customers to routes and for the sequencing of each route. The actual evaluation of the target function is achieved by solving the resulting buffer allocation model to optimality for each route separately; this sub-problem is a linear programming problem. In the terminology of **?**, our hybrid algorithm is collaborative, since there is a clear hierarchy between the two phases. Examples of earlier works that combine local search with LP are **?**, where job-machine allocation is performed via tabu search while an LP model is used for inserting buffers in between jobs. **?**, solved a VRPTW via tabu search based on the input of an LP that defines origins and destinations for full truckloads.

The parallelism between vehicle routing and production scheduling is highlighted by **?**, who study single-vehicle routing and scheduling to minimize the number of delays. Given a deadline for servicing each customer, the objective is to minimize the number of late deliveries. The problem is equivalent to single-machine scheduling with sequence-dependent setup times to minimize the number of tardy jobs. The scheduling aspect is fundamental in **?**, in the context of dynamic pickup and delivery with time windows. The authors first solve the routing component and then look into the scheduling component. Four waiting strategies are presented and assessed based on the distance along with the number of vehicles required. **?** study the dynamic dial-a-ride under various types of uncertainty. They propose several scheduling strategies for handling dynamic events, accounting for a fixed duration and overtime costs in the case of exceeding the shift length. Our problem VRP-SITW differs from the above literature in that customer demand is known in advance. Stochastic travel times in VRP are investigated in **?**, where vehicles incur penalties for exceeding a limit on the route duration. **?** examine VRPTW with stochastic travel and service times. Their model also includes overtime costs for exceeding route duration and soft time windows; the actual penalties are computed by means of simulation.

Compared to Chapter 3, this chapter does not account for time-dependent travel times. Furthermore, in this chapter we consider stochastic travel time, while in

Chapter 3 stochastic service times were considered.

The main contributions of this chapter are threefold: (1) we describe the new yet practical setting of SITW in vehicle routing; (2) we describe how a VRP with SITW and stochastic travel times can benefit from time buffers; and (3) we develop a hybrid LP / tabu search algorithm for producing high-quality solutions. Our aim is to construct a stable *a priori* plan that best copes with disruptions; in other words, a solution is generated at the start of the planning horizon and does not require further optimization during its implementation.

The remainder of the chapter is organized as follows. We provide a number of definitions and a detailed problem statement in Section **??**. Our solution procedure is described in Section **??**. The computational experiments are presented and discussed in Section **??**. Finally, in Section **??**, we highlight the main results and indicate directions for future research.

4.2. Description of VRP-SITW

Consider a set of N customers with a fleet of K identical vehicles. Each customer i has a demand q_i and is to be serviced by a single vehicle. The logistics network is represented by a complete directed graph $G = (V, A)$, with $V = \{0, 1, \ldots, N\}$ the set of vertices and A the set of directed links. The vertex 0 denotes the depot; the other vertices of V represent the customers. The non-negative weight d_{ij} associated with each arc (i, j) represents the distance from i to j. Each vehicle must start and end its route at the depot, the total demand on each route cannot exceed the vehicle capacity Q and each customer should be visited exactly once. The objective of the VRP is to construct routes that bring the total travel time of the vehicles to a minimum. The VRP-SITW entails the same elements as the VRP but with a number of additional parameters. Below, we first give a general description of the objective function (Section **??**). Subsequently, we provide we elaborate the SITW model and on the way in which stochasticity is captured (in Sections **??** and **??**, respectively).

4.2.1 Objective function

The objective function of the VRP-SITW consists of three parts. The first part is the travel cost, which captures the vehicle operating costs such as fuel costs. The second part of the objective function is a tardiness penalty, which represents the desire to respect the quoted time windows as well as possible. A 'railroad-scheduling approach' is adopted: the lower bound of the time window is the earliest starting time of the service. Arrival before the scheduled window is not penalized, since the driver cost is presumed to be fixed. However arrival after the time window leads to a penalty proportional to the tardiness. The third component of the objective function is an overtime penalty. We suppose that the drivers are paid a fixed amount for a shift with fixed duration; if this duration is exceeded then overtime penalties are due.

In the optimal solution to the VRP-SITW, the travel time will never be less than the one in the VRP since the latter does not have neither tardiness nor overtime penalties. The travel time in optimal solutions to VRP-SITW and VRPTW is in principle incomparable, since the fixed time windows are relaxed in the former but there are extra penalties in the objective. The computational experiments described in Section **??** indicate that the VRP-SITW leads to less travel time in most of the instances studied, presumably because the time windows are decision variables rather than constraints. With travel costs only, the VRP-SITW is equivalent to the VRP and is thus NP-hard.

A solution to the VRP-SITW is a set of routes $Z = \{R_1, R_2, \ldots, R_{|Z|}\}$ with $|Z| \leq K$. Each route R_r $(r = 1, \ldots, |Z|)$ is a vector $(0, i, j, \ldots, 0)$ whose components are elements of V, specifying which clients (vertices) will be visited by the vehicle following the route, and in which order. Each route begins and ends at the depot (vertex 0) and each vertex different from 0 belongs to exactly one route. We say that $i \in R_r$ if the vertex $i \in V$ is part of route $R_r \in Z$ and $(i, j) \in R_r$ if i and j are two consecutive vertices in R_r. The objective function for the VRP-SITW is then

$$F(Z) = \Omega(Z) + \sum_{R_r \in Z} \Theta(R_r), \tag{4.1}$$

with $\Omega(Z)$ the total travel cost associated with solution Z and $\Theta(R_r)$ representing the overtime and tardiness penalties of route R_r. The travel cost is defined as follows:

$$\Omega(Z) = c \sum_{R_r \in Z} \sum_{(i,j) \in R_r} d_{ij},$$

with c the cost of traveling one unit of distance. The penalties of each route are evaluated by solving a buffer allocation problem, which is described in Section **??**.

4.2.2 Self-imposed time windows

Each route R_r consists of visiting a set of n_r customers. For convenience, when referring to one specific route, we relabel the customers in ascending order: $R_r = (0, 1, 2, \ldots, n_r, n_r + 1)$, where the depot corresponds with $0 \equiv n_r + 1$. The distance $d_{i,i+1}$ between consecutive nodes i and $i + 1$ in the route is written as d_i. A *schedule* for route R_r is an $(n_r + 2)$-vector $\mathbf{s} = (s_0, s_1, \ldots, s_{n_r+1})$, specifying a departure time

s_i from each node $i \in R_r$. The shift length is the time interval $[s_s, s_e]$, implying that $s_s \leq s_0$. Each customer $i \in R_r \setminus \{0, n_r + 1\}$ has a time-window length W_i within which the arrival of the vehicle is desired. The carrier company communicates time windows to its customers based on the schedule **s**. Each node $i \in R_r$ also has a standard service time u_i, e. g., for load/unload activities. We assume that a vehicle will never leave a customer earlier than scheduled. The left bound of the time window is then $s_i - u_i$, as this constitutes an earliest starting time for the servicing operations. An illustration is provided in Figure **??**. The service times u_0 and u_{n_r+1} at the depot are set to zero.

During the realization of this baseline schedule, disruptions might occur. We examine disruptions corresponding with an increase in the travel time d_i between customers i and $i+1$. The length L_i of this delay is a random variable, which is modeled by means of discrete scenarios; a similar choice in a machine-scheduling context is made by, e. g., **?**, **?**, **?**, and **?**. Specifically, we let L_i denote the increase in d_i if i is 'disrupted', which takes place with probability p_i. The variable L_i is discrete with probability-mass function $g_i(\cdot)$, which associates non-zero probability with positive values $l_{ik} \in \Psi_i$, where Ψ_i denotes the set of disruption scenarios for d_i, so $\sum_{k \in \Psi_i} g_i(l_{ik}) = 1$. We use g_{ik} as shorthand for $g_i(l_{ik})$; the disruption lengths l_{ik} are indexed from small to large for a given i. The realization of L_i becomes known only when arc $(i, i+1)$ is traversed. The actual departure time at customer i is denoted by $s_i^a(\mathbf{s})$; this is a random variable that is dependent on the schedule **s** (in the remainder of the article, we omit the argument **s** when there is no danger of confusion). The value $s_i - u_i$ is a lower bound on the starting time of the client's service. This so-called railroad-scheduling approach implies that $s_i \leq s_i^a$, $\forall i \in R_r$, and guarantees that the actual schedule will strictly copy the baseline schedule if no disruptions occur. In effect, the scheduled times become 'release dates' for departure times s_i^a from each customer

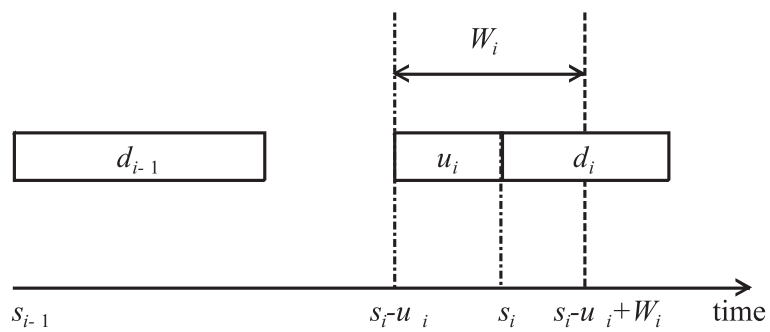

Figure 4.1 Illustration of a time window at customer i

$i \in R_r$:

$$
\begin{aligned}
s_0^a &= s_0 \\
s_i^a &= \max\{s_i; s_{i-1}^a + d_{i-1} + L_{i-1} + u_i\}, \qquad i = 1, \ldots, n_r + 1.
\end{aligned}
$$

Arrival prior to $s_i - u_i$ is not penalized. With arrival later than $s_i - u_i + W_i$, however, we associate a cost proportional to the tardiness: a non-negative integer penalty t_i is incurred per unit-time delay. The value t_{n_r+1} is the cost for arriving late at the depot at the end of the tour.

We assume that the driver receives a fixed payment for the shift, which ends at s_e. Arrival after the end of the shift incurs an overtime penalty b per time unit. We can now elaborate the penalty term $\Theta(\cdot)$ in Equation (**??**). For a given route R_r, $\Theta(R_r)$ consists of two components, namely the expected delay costs at customers and the depot on the one hand, and the expected overtime penalty on the other hand. Specifically,

$$
\Theta(R_r) = \sum_{i \in R_r \setminus \{0\}} t_i \mathbb{E}[\max\{0; s_i^a(\mathbf{s}) - (s_i - u_i + W_i)\}] + b\mathbb{E}[\max\{0; s_{n_r+1}^a - s_e\}], \quad (4.2)
$$

with $\mathbb{E}[\cdot]$ the expectation operator (note that $\mathbf{s}$ is actually also a parameter to $\Theta(\cdot)$). In the following subsection, we outline the disruption model in detail.

4.2.3 Modeling disruptions

When the durations are independent, little less is possible for objective-function evaluation than to consider all $\prod_{i \in R_r \setminus \{n_r+1\}}(|\Psi_i| + 1)$ possible combinations of duration disruptions. This was the motivation in a scheduling context in **????** to develop a model that considers only the main effects of the separate disruption of each of the individual jobs rather than all possible disruption interactions. Computational results in the aforementioned scheduling applications show that the resulting model is quite robust to variations in the actual number of disrupted jobs. In the context of time-dependent VRP with service disruptions, Chapter 3 also focus on the effects of single disruptions. We make a similar assumption in this chapter: our model assumes that in each route, exactly one leg suffers a disruption from its baseline duration. The underlying practical motivation is that we should only optimize for one 'inconvenience' per day, as it would be very difficult to protect from multiple disruptions at multiple places at multiple times. The resulting restricted model is useful when disruptions are sparse and spread over time so that the number of interactions is limited.

For a given route R_r we distinguish between two situations: 1) no leg in R_r is disturbed and 2) a single leg is disturbed in R_r. Let ζ denote the overtime for R_r when no leg is disturbed (tardiness penalties are irrelevant if no leg is disturbed).

$\Theta(R_r)$ consists of two components, namely the expected delay costs at customers and the depot on the one hand, and the expected overtime penalty on the other hand. Specifically, for a given route R_r, let p_i represent the probability that d_i is the unique disrupted value. Under the one-disruption assumption and $s_{i-1} + d_{i-1} + u_i \leq s_i$ for all $i > 0$, the relevant penalty term in (**??**) can be written as

$$\Theta(R_r) = \sum_{i=0}^{n_r} \sum_{j=i+1}^{n_r+1} \sum_{k=1}^{|\Psi_i|} p_i g_{ik} t_j \Delta_{ijk} + b \sum_{i=0}^{n_r} \sum_{k=1}^{|\Psi_i|} p_i g_{ik} \Lambda_{ik} + b\big(1 - \sum_{i=0}^{n_r} p_i\big)\zeta,$$

In this expression,

$$\Delta_{ijk} = \max\left\{0 \quad ; \quad s_i + d_i + l_{ik} + \sum_{m=i+1}^{j-1} (u_m + d_m) - s_j + u_j - W_j\right\},$$
$$i \in R_r \setminus \{n_r + 1\}; j \in R_r \setminus \{0\}; i < j; k \in \Psi_i$$

and

$$\Lambda_{ik} = \max\{0 \quad ; \quad s_{n_r+1} + \Delta_{i,n_r+1,k} - s_e\}, \qquad i \in R_r \setminus \{n_r + 1\}; k \in \Psi_i$$

and

$$\zeta = \max\{0 \quad ; \quad s_{n_r+1} - s_e\},$$

The variable Δ_{ijk} represents the tardiness at client j due to a disruption according to scenario k of d_i, which is equal to zero or to the disruption length of i minus the buffer size in place between the customers i and j, whichever is larger. The term $\sum_{m=i+1}^{j-1} (u_m + d_m)$ is the service time and travel time for the customers between i and j. Similarly, Λ_{ik} is the overtime resulting from a disruption at customer i by scenario k. The overtime is zero in case of arrival at the depot before the shift end s_e, and equal to the realized arrival time minus s_e otherwise. The probability that a route is not disturbed is $(1 - \sum_{i=0}^{n_r} p_i)$. Thus, ζ is zero in case of arrival at the depot before the shift end.

4.3. A hybrid solution procedure

Our solution method for the VRP-SITW proceeds in two stages: first routing and then scheduling. The assignment of customers to vehicles and the sequencing of customers are done in stage 1; this stage uses tabu search. Iteratively, the routes generated by the tabu search are then scheduled in the second stage, where we use linear programming to solve the sub-problem to optimality under the one-disruption assumption. We say that our solution procedure is 'hybrid' due to the combined use of a meta-heuristic and an exact optimization routine. Below, we first describe the lower-level scheduling problem in Section **??**, followed by the tabu search procedure (Section **??**).

4.3.1 Scheduling and buffer insertion

For a given route R_r, the linear program below produces an optimal schedule, conditional on exactly one leg being disrupted. Buffer sizes are implicit from the resulting schedule.

$$\Theta(R_r) = \min \qquad \sum_{i=0}^{n_r} \sum_{j=i+1}^{n_r+1} \sum_{k=1}^{|\Psi_i|} p_i g_{ik} t_j \Delta_{ijk} + b \sum_{i=0}^{n_r} \sum_{k=1}^{|\Psi_i|} p_i g_{ik} \Lambda_{ik} + b\big(1 - \sum_{i=0}^{n_r} p_i\big)\zeta$$

subject to

$$s_{i-1} + d_{i-1} + u_i \leq s_i \qquad i \in R_r \setminus \{0\} \tag{4.3}$$

$$s_0 \geq s_s \tag{4.4}$$

$$s_i + d_i + l_{ik} + \sum_{m=i+1}^{j-1} (u_m + d_m) \leq s_j - u_j + W_j + \Delta_{ijk}$$

$$i \in R_r \setminus \{n_r + 1\}; j \in R_r \setminus \{0\}; i < j; k \in \Psi_i \tag{4.5}$$

$$s_{n_r+1} + \Delta_{i,n_r+1,k} - s_e \leq \Lambda_{ik} \qquad i \in R_r \setminus \{n_r + 1\}; k \in \Psi_i \tag{4.6}$$

$$\zeta \geq s_{n_r+1} - s_e \tag{4.7}$$

$$\text{all } \Delta_{ijk} \geq 0; \text{ all } s_i \geq 0; \text{ all } \Lambda_{ik} \geq 0; \ \zeta \geq 0 \tag{4.8}$$

Constraints (**??**) can be viewed as precedence constraints: the scheduled departure time s_i from customer i is at least equal to the departure time of its predecessor s_{i-1} augmented with the distance d_{i-1} and the service time u_i. This implies that the buffer between customers $i-1$ and i is $s_i - [s_{i-1} + d_{i-1} + u_i]$. Constraint (**??**) ensures that the scheduled departure time from the depot does not precede the shift's start time s_s. Constraints (**??**),(**??**) and (**??**) determine the delay terms Δ_{ijk}, Λ_{ik} and ζ respectively, as described in Section **??**.

4.3.2 Tabu search for the VRP-SITW

Tabu search has been widely used for solving the VRP, e. g., **???**. Furthermore, it has been extensively used to solve VRPTW as well, examples can be found in **??**. Thus, adopting the tabu search heuristic comes as a natural choice also for the VRP-SITW. Our tabu search procedure generates a set of routes that still need to be scheduled using the lower-level LP described in Section **??**. The procedure iteratively scans the members of a neighborhood of the current solution to evaluate possible improvements in the objective function. Due to our bi-level approach, the evaluation of each neighborhood solution requires a separate LP run, which, if performed to optimality, would require enormous computation times. We have therefore opted for approximating the optimal overtime and tardiness penalties heuristically, which will guide the tabu search in selecting the best move in its current neighborhood. Once a move is selected, its exact target function is computed by invoking the LP model for the changed route or routes, leading to an optimal schedule.

The overall procedure is described in pseudo-code as Algorithm **??**. We adopt three different criteria C_1, C_2 and C_3 for choosing a move; these will be described in detail below. The tabu search procedure is run consecutively with each of the three criteria. The initial solution Z_0 is the output of the nearest neighbor heuristic for each of the three criteria. Feeding the best-found solution of C_1 into the run for C_2 and for C_2 into C_3 has been tested, together with many variations of the order of the three criteria, but this did not lead to better results. For each customer $i \in V$, we construct 2-opt* **?** and Or-opt **?** neighborhoods for the η nodes closest to i. A chosen move is declared tabu for the next κ iterations. The process iterates until a maximum number of non-improving moves is reached.

In line with **?**, diversification of the search is achieved by allowing demand-infeasible solutions (i. e., routes with total demand exceeding the vehicle capacity). Such infeasible solutions are penalized in proportion to their capacity violation by means of the following composite objective function, which replaces $\Omega(Z)$:

$$\Omega_2(Z) = \Omega(Z) + w \sum_{R_r \in Z} \left[\Big(\sum_{i \in R} q_i \Big) - Q \right]^+ . \tag{4.9}$$

In Equation (**??**) each unit of excess demand is penalized by a factor w. This excess penalty w is decreased by multiplication with a factor ν after ϕ consecutive feasible moves. Similarly, w is increased (multiplied by factor ν^{-1}) after ϕ infeasible iterations.

Below, we describe the three criteria that allow avoiding the use of the LP model for each candidate solution and lead to computationally efficient move selection procedures.

C_1 **- distance based** This heuristic is based purely on minimizing the modified travel costs $\Omega_2(\cdot)$, i. e., it does not take into account the time windows and their associated penalties, nor does it consider overtime. Thus, C_1 is similar to the criteria used in local search for the VRP. Let Z' be a neighbor of the current solution Z and define $\Delta_1(Z') = \Omega_2(Z) - \Omega_2(Z')$. The chosen move is

Algorithm 2 Global algorithmic structure

1: construct initial solution Z_0 and compute $F(Z_0)$
2: **for** $\xi = 1$ to 3 **do**
3: set $Z = Z_0$ and $F(Z) = F(Z_0)$
4: generate the neighborhood of Z
5: evaluate all neighbors on criterion C_ξ and retain the best non-tabu move as new solution Z
6: evaluate $F(Z)$ and update the tabu list to include Z
7: **if** Z is feasible and is better than the current best solution **then**
8: update the best feasible solution for C_ξ to Z
9: **end if**
10: update excess demand penalty
11: **if** no improvement in $\eta_{\max}$ iterations **then**
12: store best solution for C_ξ
13: **else**
14: go to step 4
15: **end if**
16: **end for**
17: return the best solution from $\xi = 1, 2$ and 3

one that is not tabu and maximizes $\Delta_1(\cdot)$.

C_2 **- distance based and marginal penalties** This measure adds to C_1 an assessment of the penalty component $\sum_{R_r \in Z}(\Theta(R_r))$. For given Z, the marginal penalty of route R_r is $\frac{\Theta(R_r)}{n_r+1}$. Consider a move involving two routes R_1 and R_2, leading to solution Z' . Let n_1 and n_2 be the number of nodes visited by routes R_1 and R_2, respectively, in the current solution Z, and n'_1 and n'_2 the number of nodes visited by routes R_1 and R_2 in the new solution Z'. C_2 picks the move that is not tabu and maximizes the following expression:

$$\Delta_2(Z') = [\Omega_2(Z) - \Omega_2(Z')] + \rho\Big[\Theta(R_1) + \Theta(R_2) - \frac{\Theta(R_1)}{n_1+1}(n'_1+1) - \frac{\Theta(R_2)}{n_2+1}(n'_2+1)\Big].$$

The logic behind this evaluation is based on the observation that penalties increase with the number of customers in the route. Decreasing the number of customers in a route with a large penalty value is likely to decrease the total objective value associated with the route.

C_3 **- distance and buffer based** As mentioned in Section **??**, the buffer size between customers i and $i+1$ is $b(i) = s_{i+1} - [s_i + d_i + u_{i+1}]$. C_3 favors moves with small buffers. Each buffer unit is penalized by γ. For each candidate solution Z' involving a move between customer i and customer j, we compute the following quantity:

$$\Delta_3(Z') = [\Omega_2(Z) - \Omega_2(Z')] - \gamma[b(i) + b(j)].$$

We chose a move that is not tabu and that maximizes $\Delta_3(\cdot)$. The reasoning involved in this move selection process is the following: improvements in travel times are more likely to also decrease the penalties when the buffers are small.

Different aspects of the problem are tackled by each criterion. The impact of a move on the travel time $\Omega_2(Z)$ is efficiently computed. The accurate impact of a move on the penalty component $\sum_{R_r \in Z} \Theta(R_r)$ of the target function, on the other hand, requires evaluation of the SITW model for the affected route or routes. Criteria 2 and 3 attempt to assess moves based on the penalty values of the current solution rather than via the LP model. We note that C_2 is equivalent to C_1 for moves involving a single route, which can occur only with Or-opt moves.

4.4. Computational experiments

We have run a number of experiments to assess the computational performance of our algorithm and to compare the outcomes of the VRP-SITW with both the results of the VRP and of the VRPTW. Throughout this section, the travel cost c in $\Omega(Z)$ is set to one, thus we use the terms distance and travel time interchangeably. For an instance with N nodes, for each customer the $\eta = \lceil 0.3N \rceil$ closest customers are candidates for a move. The tenure size κ is set to 20. The infeasibility penalty w equals 12, with $\phi = 5$ and $\nu = \frac{3}{4}$. The penalties associated with C_2 and C_3 are chosen as $\rho = 1$ and $\gamma = 0.1$, respectively. The overtime penalty β takes the value 2. The probability p_i is set to one over the total number of legs in a solution. Given a solution with k vehicles, where $k \leq K$, $p_i = \frac{1}{N+k}$. Hence, the probability of disruption is identical for all the legs in the solution.

We consider four disruption scenarios for each leg: $|\Psi_i| = 4$. The probabilities of disruption are also the same for each leg i, namely $g_{i1} = 0.5$, $g_{i2} = 0.3$, $g_{i3} = 0.1$ and $g_{i4} = 0.1$. Finally, the disruption lengths between customers i and j are assumed proportional to the baseline duration d_{ij}, namely $l_{i1} = 0.1d_{ij}$, $l_{i2} = 0.2d_{ij}$, $l_{i3} = 0.5d_{ij}$ and $l_{i3} = d_{ij}$.

All experiments are performed on a Intel(R)Core Duo with 2.40 GHz and 2 GB of RAM. The implementation is coded in C++. The LP instances are solved by embedding Gurobi Optimizer 2.0.2, which uses the simplex algorithm. The reported run times are in seconds. We have adopted two datasets from the literature. The first dataset contains a number of VRP instances from **?**. We work with 27 VRP instances, with the number of customers ranging from 31 to 79. The vehicle capacity Q is 100 units. The baseline service time u_i for each customer i is set to 10 minutes. The shift start time and end time s_s and s_e are chosen as zero and 200, respectively. The window length W_i equals 60 for all i. The second dataset contains VRPTW instances and stems from **?**. We consider 29 instances with 100 customers (sets R1 (random), C1 (clustered) and RC1 (random and clustered)). The baseline service times u_i and window sizes W_i are given. The opening hours of the depot are used to determine the shift's starting time s_s and ending time s_e. The vehicle capacity Q is 200 units.

Below, we first conduct some experiments related to move selection and tardiness choices (in Section **??** and **??**, respectively), followed by comparisons with VRP (Section **??**) and with VRPTW (Section **??**).

4.4.1 Move selection

Table **??** shows the results of implementations for the Augerat instances in which only one of the three criteria C_1, C_2 and C_3 is used during the optimization; the tardiness penalty $t_i = 5$ for all arcs. The left side of the table displays the target function value $F(Z)$ attained. The right side of the table exhibits the run time for each of the three measures. We observe that C_3 outperforms the other two critera in 15 out of the 27 instances, while C_1 and C_2 do so in 7 and 5 instances, respectively. On average, C_1 requires less runtime than C_2 and C_3. The average run time over all heuristics is 17.3 minutes. Since we are working in an *a priori* setting, these running times are acceptable.

Table **??** contains similar results for the Solomon instances. The computation times are larger than those for the first dataset. This is partly due to a greater number of customers, but more importantly the number of customers per route is also larger than before. Thus, the LP subroutine will consume considerably more time. We note that we obtain identical results for some of the instances, which is due to the fact that the time window constraints in these VRPTW instances are now relaxed, and some of instances have the same time window lengths and customer locations. In line with Table **??**, the three move selection criteria differ in performance. C_2 performs best in 23 out of the 27 instances, while this occurs for C_1 and C_3 in 2 and 4 instances, respectively.

	Objective value			CPU time			
	C_1	C_2	C_3	C_1	C_2	C_3	Total
32 k5	955.4	1038.2	957.2	734	103	1568	2405
33 k5	744.8	724.1	716.6	78	121	166	365
33 k6	801.1	798.7	791.0	213	151	177	541
34 k5	867.9	876.1	852.7	335	135	374	844
36 k5	958.0	990.1	950.5	552	222	438	1212
37 k5	765.6	811.8	798.6	338	394	210	942
37 k6	1071.1	1069.0	1080.5	112	158	148	418
38 k5	822.6	832.5	823.4	361	299	227	887
39 k5	1013.1	957.5	995.9	200	302	289	791
39 k6	963.0	956.1	952.7	184	130	151	465
44 k6	1102.9	1057.7	1054.7	128	124	175	427
45 k6	1078.0	2685.8	1096.4	1117	71	1142	2330
45 k7	1294.4	1281.4	1302.8	80	86	82	248
46 k7	1072.5	1059.0	1008.7	99	221	401	721
48 k7	1256.3	1243.1	1247.2	169	230	224	623
53 k7	1185.3	1194.7	1165.3	192	1046	376	1614
54 k7	1293.7	1396.7	1335.5	253	446	320	1019
55 k9	1158.7	1137.4	1132.2	340	212	255	807
60 k9	1509.2	1489.4	1473.8	108	112	177	397
61 k9	1197.9	1239.7	1177.3	225	224	214	663
62 k8	1509.7	1516.0	1499.5	295	893	386	1574
63 k10	1556.2	1411.1	1493.0	157	607	292	1056
63 k9	1834.5	1897.8	1840.8	343	317	712	1372
64 k9	1658.5	1626.5	1587.8	202	431	521	1154
65 k9	1319.7	1307.3	1293.2	137	1249	115	1501
69 k9	1254.5	1276.8	1291.3	616	452	552	1620
80 k10	2095.0	2057.7	2046.5	399	1002	693	2094
Average				295	361	385	1040

Table 4.1 Comparison of the three move selection criteria for the Augerat instances

4.4.2 Tardiness penalty choices

In order to evaluate the effects of varying delay penalty costs t_i, we have conducted experiments under four different cost settings, which are subsequently referred to as 'P5', 'P10', 'Prob' and '1.3dist'. In P5, we choose $t_i = 5, \forall i \in V \setminus \{0\}$ (which was the choice also in Section **??**), while P10 corresponds to $t_i = 10$. Under setting Prop, the delay cost for each customer equals the quantity ordered, so $t_i = q_i, \forall i \in V \setminus \{0\}$, which represents a situation where the delay penalty is proportional to the demand. The final experimental setting, denoted by 1.3dist, puts t_i equal to 5 for all customers, similarly to P5, but all distances are now increased by 30%. In this way, there is less slack time available, leading to less buffer time to be allocated and resulting in tighter instances.

	Objective value			CPU time			
	C_1	C_2	C_3	C_1	C_2	C_3	Total
R101	918.6	905.7	922.4	1773	3388	775	5936
R102	918.2	922.5	922.1	1724	1860	769	4353
R103	918.2	922.5	922.1	1734	1865	774	4373
R104	917.2	917.0	920.5	1737	3445	771	5953
R105	917.0	908.8	920.1	1722	2412	767	4901
R106	917.0	908.8	920.1	1743	2384	761	4888
R107	917.0	908.8	920.1	1752	2362	773	4887
R108	917.0	908.8	920.1	1765	2392	764	4921
R109	917.0	908.8	920.1	1728	2375	767	4870
R110	917.0	908.8	920.1	1730	2240	761	4731
R111	917.0	908.8	920.1	1717	2229	768	4714
R112	917.0	908.8	920.1	1743	2240	761	4744
C101	834.7	834.6	859.2	805	3209	1315	5329
C102	834.7	834.6	859.2	802	3181	1328	5311
C103	834.7	834.6	859.2	799	3210	1317	5326
C104	834.7	834.6	859.2	807	3271	1324	5402
C105	834.7	834.6	859.2	796	3308	1317	5421
C106	834.7	834.6	859.2	799	3296	1316	5411
C107	834.7	834.6	859.2	798	3211	1327	5336
C108	834.7	834.6	859.2	803	3187	1319	5309
C109	834.7	834.6	859.2	792	3198	1327	5317
RC101	1024.5	1013.2	1022.6	1198	1318	1071	3587
RC102	1024.5	1013.2	1022.6	1196	1318	1075	3589
RC103	1024.5	1013.4	1022.6	1195	1740	1121	4056
RC104	1024.5	1042.0	1022.6	1201	1026	1137	3364
RC105	1025.0	1013.6	1023.2	1195	1318	1093	3606
RC106	1024.5	1042.0	1022.6	1189	1024	1083	3296
RC107	1024.5	1042.0	1022.6	1209	1021	1090	3320
RC108	1024.5	1042.0	1022.6	1191	1023	1083	3297
Average				1298	2347	1029	4674

Table 4.2 Comparison of the three move selection criteria for the Solomon instances

Table **??** summarizes the results for the four experimental settings after running the full tabu search procedure (with the three criteria combined). The left side of the table shows the achieved target function values. $M(C_i)$ denotes the number of times (out of 27) that criterion C_i produces the best result; these values are presented in the last three lines of the table. C_3 performs best in more instances in all 4 experimental setting. The best result for C_3 is in P5. In total C_2 and C_3 perform best in 30 and 26 instances, respectively, when considering the all four experimental settings. The fact that C_3 takes accounts for buffer between customers its superior performance.

On average, the objective values for P10 are only 0.7% higher than P5. This means that even doubling the customer delay penalty does not affect the final objective value to a large extent. With varying penalties, as in the Prop setting, the values are not

Set	Objective				Penalty ratio			
	P5	P10	Prop	1.3dist	P5	P10	Prop	1.3dist
32 k5	955.4	961.7	956.6	1290.8	16.6%	17.1%	16.7%	18.5%
33 k5	716.6	716.9	716.5	998.7	6.3%	5.0%	6.3%	12.4%
33 k6	791.0	796.8	797.6	1066.9	5.0%	5.7%	5.8%	8.9%
34 k5	852.7	857.6	857.2	1190.6	7.0%	7.6%	7.6%	13.4%
36 k5	950.5	960.3	957.8	1285.6	13.5%	14.4%	14.2%	17.8%
37 k5	765.6	766.8	766.4	1101.0	10.9%	11.0%	11.0%	17.0%
37 k6	1069.0	1079.4	1079.3	1457.3	9.2%	9.6%	9.6%	13.3%
38 k5	822.6	824.3	824.0	1162.0	7.6%	7.8%	7.8%	15.2%
39 k5	957.5	971.6	969.4	1283.2	11.2%	11.6%	11.4%	15.3%
39 k6	952.7	957.8	953.4	1295.1	10.5%	11.0%	8.7%	16.0%
44 k6	1054.7	1059.6	1059.2	1489.6	8.8%	9.3%	9.2%	13.4%
45 k6	1078.0	1081.3	1066.1	1469.0	8.4%	8.6%	8.2%	12.7%
45 k7	1281.4	1298.3	1277.7	1713.5	7.7%	8.9%	8.5%	11.3%
46 k7	1008.7	1007.5	1009.4	1371.5	7.0%	6.9%	7.1%	10.6%
48 k7	1243.1	1244.1	1231.5	1662.6	10.2%	9.0%	8.2%	11.8%
53 k7	1165.3	1168.2	1167.1	1542.7	7.4%	7.6%	7.5%	11.7%
54 k7	1293.7	1302.2	1302.5	1799.6	7.6%	8.2%	8.2%	12.3%
55 k9	1132.2	1135.5	1136.8	1506.1	2.7%	2.9%	3.0%	5.1%
60 k9	1473.8	1482.7	1485.1	1980.8	5.5%	5.1%	6.2%	8.0%
61 k9	1177.3	1178.6	1178.5	1651.2	3.4%	3.6%	3.5%	6.6%
62 k8	1499.5	1505.7	1486.1	1986.3	9.4%	9.7%	8.7%	11.9%
63 k10	1411.1	1500.0	1501.2	1914.9	4.0%	4.1%	4.2%	6.8%
63 k9	1834.5	1847.7	1844.1	2472.9	8.5%	9.2%	9.0%	11.3%
64 k9	1587.8	1598.7	1597.1	2166.2	8.5%	9.1%	9.0%	11.3%
65 k9	1293.2	1295.3	1293.7	1720.5	3.0%	3.2%	3.1%	6.4%
69 k9	1254.5	1256.6	1256.7	1643.9	4.0%	4.1%	4.1%	6.4%
80 k10	2046.5	2061.4	2042.8	2756.5	9.9%	10.6%	9.1%	11.7%
Average penalty %					7.9%	8.2%	8.0%	11.7%
$M(C_1)$	7	8	7	8				
$M(C_2)$	5	5	7	9				
$M(C_3)$	15	14	13	10				

Table 4.3 Results for the Augerat instances with four different penalty settings

dramatically different either. For the case of 1.3dist, the average objective increase is 36.1% compared to P5, while the distances are raised by only 30%. This difference can be explained by the fact that when distances rise, there is less buffer time to be allocated and the solutions are more prone to suffer overtime and delay penalties.

The right part of Table **??** shows the 'penalty ratio', the proportion

$$\sum_{R_r \in Z} \frac{\Theta(R_r)}{F(Z)}$$

of the total objective that corresponds to penalties. The average over all four experimental sets is 9.0%. The lowest ratios are achieved for P5 and Prop, followed by P10, and the ratios for 1.3dist are by far the largest. We conclude that an increase in the distances has a substantial impact on the delay penalties.

4.4.3 VRP-SITW versus VRP

The addition of SITW to the VRP can be expected to affect the distance traveled and the number of vehicles used. To assess the effect, we compare the total distance in VRP-SITW with the optimal VRP solutions (taken from **?**). The details are provided in Table **??**. On average for P5 and P10 the additional distance is 3.3% and 3.7%, respectively, which shows that, at least as far as distance minimization is concerned, our heuristic solutions are rather close to optimality. For 1.3dist the VRP distances are scaled by 1.3. The distance increase is not substantial for any of the sets.

	Increase in distance			
	P5	P10	Prop	1.3 dist
32 k5	101.1%	101.1%	102.7%	101.1%
33 k5	101.3%	102.7%	101.6%	101.3%
33 k6	101.2%	101.2%	100.7%	101.2%
34 k5	101.5%	101.4%	101.5%	101.4%
36 k5	102.5%	102.5%	101.3%	102.5%
37 k5	101.4%	101.4%	104.5%	101.4%
37 k6	102.0%	102.5%	102.1%	102.5%
38 k5	103.5%	103.5%	103.3%	103.5%
39 k5	102.6%	103.7%	100.9%	103.7%
39 k6	102.3%	102.3%	100.5%	104.5%
44 k6	102.4%	102.4%	105.6%	102.4%
45 k6	104.6%	104.6%	104.5%	103.6%
45 k7	103.1%	103.1%	101.9%	101.9%
46 k7	102.1%	102.1%	102.7%	102.1%
48 k7	103.9%	105.4%	105.0%	105.2%
53 k7	106.5%	106.5%	103.4%	106.5%
54 k7	102.0%	102.0%	103.6%	102.0%
55 k9	102.6%	102.6%	102.4%	102.6%
60 k9	102.7%	103.8%	103.4%	102.7%
61 k9	109.4%	109.4%	114.2%	109.4%
62 k8	105.0%	105.0%	104.0%	104.9%
63 k10	103.1%	109.5%	104.5%	109.5%
63 k9	103.4%	103.4%	104.0%	103.4%
64 k9	103.7%	103.7%	105.6%	103.7%
65 k9	106.1%	106.1%	104.8%	106.1%
69 k9	103.3%	103.3%	101.5%	103.3%
80 k10	104.4%	104.4%	106.0%	105.1%
Average	103.3%	103.7%	103.5%	103.6%

Table 4.4 Comparison of VRP-SITW with optimal VRP solutions for the Augerat instances

4.4.4 VRP-SITW versus VRPTW

The goal of this section is to evaluate the benefits of the flexibility in setting time windows compared to exogenously predetermined time windows. To this aim, we work with 29 VRPTW instances from **?**. We compare the results of the VRP-SITW with the best-known solutions for the Solomon instances as reported in **?**.

	T_F	T_S/T_F	K_F	$K_F - K_S$	$\sum_{R_r \in Z} \Theta(R_r)/F(Z)$
R101	1637.7	52.0%	20	12	6.3%
R102	1466.6	59.9%	18	10	4.5%
R103	1208.7	72.7%	14	6	4.5%
R104	971.5	89.7%	11	3	5.2%
R105	1355.3	63.5%	15	7	5.6%
R106	1251.98	68.7%	12	4	5.6%
R107	1064.6	80.8%	11	3	5.6%
R108	960.88	89.5%	9	1	5.6%
R109	1146.9	75.0%	13	5	5.6%
R110	1068	80.5%	12	4	5.6%
R111	1048.7	82.0%	12	4	5.6%
R112	982.14	87.6%	9	1	5.6%
C101	827.3	100.9%	10	0	0.0%
C102	827.3	100.9%	10	0	0.0%
C103	826.3	101.0%	10	0	0.0%
C104	822.9	101.4%	10	0	0.0%
C105	827.3	100.9%	10	0	0.0%
C106	827.3	100.9%	10	0	0.0%
C107	827.3	100.9%	10	0	0.0%
C108	827.3	100.9%	10	0	0.0%
C109	827.3	100.9%	10	0	0.0%
RC101	1619.8	61.9%	15	6	1.1%
RC102	1457.4	68.8%	14	5	1.1%
RC103	1258	79.4%	13	4	1.4%
RC104	1261.67	79.6%	11	2	1.7%
RC105	1513.7	66.2%	15	6	1.1%
RC106	1424.73	70.5%	11	2	1.7%
RC107	1207.8	83.2%	12	3	1.7%
Average		82.9%			2.7%

Table 4.5 Comparison of VRP-SITW with the best known VRPTW solutions for the Solomon instances

Table **??** reports the results. For practical purposes we denote the travel time, which is equivalent to the distance, by T_F and T_S for the VRPTW (which has fixed time windows) and the VRP-SITW, respectively. The number of vehicles required in the VRPTW is denoted by K_F while the number of vehicles used by the VRP-SITW solution is denoted by K_S. The third column in Table **??** gives the ratio of the total travel times between both solutions. We observe that the VRP-SITW substantially reduces the travel time for instances with tight time windows such as those in the R1 and RC1 sets. Sets C1, on the other hand, achieves zero penalty values, which can be read from the last column of the table. We conclude that these instances have quite unrestrictive time windows and exhibit a behavior similar to the VRP instances studied in Section **??**. In general the penalty component $\sum_{R_r \in Z} \Theta(R_r)$ comprised at most 6.3 % of the total objective value.

The fifth column of Table **??** displays the number of vehicles saved in VRP-SITW compared to VRPTW. A substantial reduction in the required number of vehicles is observed in the R1 and RC1 sets. In set C1, however, no such reduction is achieved. We conclude that those instances that are eligible for substantial reductions in travel times present the same behavior with respect to the number of vehicles.

We ran the VRP-SITW given the initial number of vehicles set to K_F, as reported for the VRPTW sets. However, no additional vehicles were utilized, i. e., K_S values were similar to the ones reported in Table **??**. Additional vehicles may improve upon $\sum_{R_r \in Z} \Theta(R_r)$ yet may result in additional travel times. The fact that $\sum_{R_r \in Z} \Theta(R_r)$ values are relatively small, 2.7% on average, explains the outcome of not making use of additional vehicles. Thus, an additional vehicle maybe used if customer delay is valued at a higher cost or when solutions are subject to larger disruptions.

4.5. Conclusions

In this chapter, we have analyzed the situation of carrier companies that face the problem of making routing decisions combined with the quotation of arrival times to their customers; we have referred to this setting by the term 'Self-Imposed Time Windows' (SITW). In the context of vehicle routing, the resulting VRP-SITW extends the VRP by the incorporation of customer specific service aspects, reflected in the carrier company's ability to uphold the time windows once they have been quoted in a stochastic environment. In comparison with the VRP with exogenous time windows (VRPTW), the customer service requirement is somewhat relaxed, in that the service provider has ex ante flexibility in choosing a convenient time interval that will be quoted.

Our solution approach is a hybrid algorithm that is comprised of two main components: routing and scheduling. The routing component is handled via a tabu search procedure, while scheduling is performed by solving an LP model that implicitly inserts buffers into each route's schedule.

We have compared the VRP to VRP-SITW under different choices for penalty structures and distances. In our tests, the results indicate that the VRP-SITW requires an average increase of some 3.5% in distance. Further research might focus on the impact of additional vehicles on the penalties.

Contrary to the VRP, we observe substantial differences on comparing VRPTW to VRP-SITW. In most cases, the VRP-SITW requires significantly less distance and uses far less vehicles. Clearly, the VRP-SITW benefits greatly from its flexibility in setting the time windows. In our opinion, there is important potential in conducting an in-depth study of various flexibility levels in choosing delivery windows. Such a study can be beneficial, for instance, when negotiating service contracts. Another extension might look into the setting where only a subset of customers have fixed time windows. Furthermore, given some alterations the proposed model can also accommodate driving breaks, by using the buffers for the breaks. The proposed model establishes an *a priori* plan for a static environment. A major extension of the model might incorporate the quotation of time windows for dynamically arriving orders.

Investigating the added value of additional vehicles is a valid extension to the VRP-SITW. Generally, additional vehicles will improve the ability to uphold SITW. However, additional vehicles may increase travel times. This trade-off may be studied along with attributing a fixed cost per vehicle.

Chapter 5

Consistent Vehicle Routing with Stochastic Customers

5.1. Introduction

The Vehicle Routing Problem (VRP) has been extensively studied for more than fifty years (for a comprehensive literature survey see **?**). In its core lies the objective of minimizing operational cost, traditionally modeled by the sum of distances or travel times. Incorporating customer service measures or restrictions to the VRP has been an evolving research direction. One service aspect that has been recently introduced is the notion of consistent service. In this context, **?** defined consistency as having the same driver visiting the same customers at roughly the same time on each day that these customers require service. This definition stems from the needs expressed by United Parcel Service (UPS), as providing consistent service is a tangible feature in their operations.

Developing consistent routes is beneficial from an operational perspective, since it increases driver familiarity with the regions he serves on a daily basis. From a customer service standpoint, consistent service enables building real bonds with customers. In the case of UPS, drivers have managed to increase sales by 60 million packages per year **?**. Furthermore, consistency preserves stability in the routes which translates to less alterations in schedules, when compared to carrying out a full optimization upon realization of customer demand.

Several ways have been proposed in the literature to coupe with consistency. **?** examined assigning drivers to territories. In the context of the periodic vehicle routing problem, **?** treated three consistency measures. Driver consistency was defined as minimizing the number of different drivers serving each customer. Customer familiarity was treated as maximizing the number of times that a unique driver visits a customer. Finally, region familiarity was also considered and measured by maximizing the number of times that a driver repeatedly visits a region.

? treated driver consistency as an objective by maximizing the number of customers who are serviced by the same driver, and who are in the vicinity of his master route. Our model resembles most the one introduced by **?**, where customers appear with a certain probability and should be serviced by the same driver. Furthermore each customer is to be visited within a time window.

The recent research, in the field of consistent vehicle routing, has been conducted under the assumption of known customer occurrence for a number of days. Thus, solution procedures are constructed accordingly in **?** and **?**. While **?** proposed a scenario-based stochastic programming heuristic. We consider the consistent vehicle routing with stochastic customers (SConVRP), i. e., customers place orders with a given probability. This relaxes the assumption of comprehensive periodic knowledge of customer occurrence. Driver consistency is imposed by having the same driver visiting the same customers when they place an order. Furthermore, we impose temporal consistency by which customers are visited at the same time interval, as much as possible, if they place an order. This is done by using Self-Imposed Time windows (SITW), which in essence means that the carrier company decides on a time window for visiting a customer and respects it as much as possible. The aim of this paper is to define the SConVRP and propose an exact solution approach.

Driver consistency coupled with temporal consistency dictate using an *a priori* solution approach. Hence, we adopt the 0-1 integer L-shaped algorithm (introduced by **?**) for solving the SConVRP. The vehicle assignment to customers and customer service time windows are set in stage 1. In stage 2 the customer occurrences are realized, the sequence set in stage 1 is executed, and penalties and travel times are computed. The penalties are computed based on the deviations from the customers time windows. Consequently, the recourse is the sum of penalties associated with early and late arrivals as well as the total travelling time.

The rest of the paper is organized as follows. The model formulation is described in Section **??**. Section **??** details the implementation of the 0-1 integer L-shaped algorithm. In Section **??** computational experiments are presented and directions for future research are indicated. Finally, Section **??** highlights the main results.

5.2. Model

We model the SConVRP using a two-index stochastic VRP formulation. Let $G = (V, E)$ define an undirected graph, where $V = \{v_1, \ldots, v_n\}$ is a set of vertices and $E = \{(v_i, v_j) : v_i, v_j \in V, i < j\}$ is a set of edges. Given the elements of set V, let vertex v_1 correspond to a depot, from which a fleet of m vehicles are used to visit a set of clients represented by vertices $v_2, \ldots, v_n$. Associated with each edge $(v_i, v_j) \in E$, let c_{ij} denote the travel time of visiting client j immediately after visiting client i. We assume that these travel times are symmetric and that the triangle inequality holds for all travel times. Let us also recall that client occurrences are stochastic in the present model. Therefore, we consider that each client i (for $2 \leq i \leq n$), requires a visit according to a Bernoulli distribution with value p_i ($0 \leq p_i \leq 1$), defined here as the probability of occurrence associated with client i. These random variables are considered to be independent from one another. We define Ω as the set of possible scenarios (or outcomes) associated with the occurrences for all customers. We also denote by $\mathcal{A}$ a collection of random events defined according to Ω.

In our formulation, the *a priori* plan, that is established in the first stage, is made up of the following: a series of m routes (i. e., one assigned to each available vehicle) that visit all customers once and a set of self-imposed time windows quoted for each customer. These time windows, which are communicated to the customers, represent the expected arrival times of the vehicles whenever visits are requested. In the second stage, a particular scenario $\omega \in \Omega$ is observed and routes, as established in the *a priori* plan, are executed. In this case, the recourse defines a cost, which is comprised of penalties incurred for violating the established time windows as well as the observed travel times of the routes. It should be noted that customers not needing a visit are simply skipped when following the sequence defined by the routes. Driver consistency is guaranteed here, given that customers are always visited by the same vehicle assigned to them in the *a priori* plan. However, temporal consistency might not be respected given that, in the second stage, vehicles are dispatched to serve only the occurring customers. It is considered that a vehicle immediately starts its service when it arrives at a customer's location. Early or late arrivals are simply penalized according to deviations from the quoted time windows established in the first stage. In this chapter, deviations from the time windows are treated as a service performance measure. However, time windows do not change the execution of a route.

Each vehicle is assumed to have enough capacity to serve all customers, thus customer demands are irrelevant in our model. However, route duration constraints are imposed on each vehicle to account for time restrictions on working shifts performed by drivers. Therefore, in the first stage of the model, we impose that a given vehicle can travel at most λ hours per route. In addition to making the model more realistic, these restrictions also ensure that the obtained routes are balanced. As for the overall objective of the model, we define it as the minimization of the total expected cost over the given *a priori* plan (i. e., the set of m routes and the established time windows for the customers). The total cost being represented here as the sum of both the total travelling times of the vehicles and the total penalty costs, which will be clearly formulated further on in this section.

The SConVRP model differs from the VRP-SITW model, presented in Chapter 4. In the VRP-SITW model customers are fixed and travel times are stochastic, while in SConVRP model travel times are fixed and customers are stochastic. Another distinction between the two models is in the execution policy, in VRP-SITW, arriving before a customer time window implies waiting. However, in SConVRP arriving before the time window translates into a penalty. This allows for an independent treatment of customers, as we will demonstrate in the remainder of this section.

We can now define the first stage decision variables as follows: let x_{ij} be equal to one if client j is visited immediately after client i, and zero otherwise. It should be noted that variables x_{1j} can also take value two if a vehicle only visits client j on its route. As for the self-imposed time windows, they are set in the following way: let t_i define a target arrival time for customer i ($2 \leq i \leq n$). Customer i is then quoted an expected arrival time between $[t_i - w, t_i + w]$, where w is the half width of the time window.

Considering the stochastic parameters within the model (i. e., the occurrences of customers), let ξ be a random vector containing all random Bernoulli variables associated with vertices $V \setminus \{v_1\}$. For each scenario $\omega \in \Omega$ let $\xi(\omega)^\top = [\xi_2(\omega), \dots, \xi_n(\omega)]$ be the corresponding scenario vector associated with ξ, where values $\xi_i(\omega)$, for $2 \leq i \leq n$, are defined as follows:

$$\xi_i(\omega) = \begin{cases} 1 & \text{if client } i \text{ is present in } \omega \in \Omega \\ 0 & \text{otherwise} \end{cases}$$

If $T(x, \xi(\omega))$ is defined as the second stage cost, (i. e., operating the *a priori* plan given that scenario $\omega \in \Omega$ is observed in the second stage), then the SConVRP is formulated as follows:

$$\min_x Q(x) = \mathbb{E}_\xi T(x, \xi(\omega)) \tag{5.1}$$

subject to

$$\sum_{j=2}^{n} x_{1j} = 2m \tag{5.2}$$

$$\sum_{i<k} x_{ik} + \sum_{j>k} x_{kj} = 2 \qquad k = 2, \ldots, n \tag{5.3}$$

$$\sum_{i,j \in S} x_{ij} \le |S| - \left\lceil \frac{\ell(S^+)}{\lambda} \right\rceil \qquad S \subset V \setminus \{v_1\}, 2 \le |S| \le n-2 \tag{5.4}$$

$$0 \le x_{1j} \le 2 \qquad j = 2, \ldots, n \tag{5.5}$$

$$0 \le x_{ij} \le 1 \qquad 2 \le i < j \le n \tag{5.6}$$

$$x_{ij} \text{ integer} \qquad 1 \le i < j \le n \tag{5.7}$$

Similar to **?**, where $S^+ = S \cup \{1\}$, the $l(S^+)$ values are defined as follows

$$\ell(S^+) = \sum_{i,j \in S^+ \subset V} c_{ij} x_{ij}.$$

Constraints (**??**) and (**??**) are standard constraints, which define the degree of each vertex. Constraints (**??**) prohibit disconnected subtours from the depot, as well as, subtours having a total maximum length exceeding λ (i. e., the route duration restriction). As for constraints (**??**), (**??**) and (**??**), they impose both non-negativity and integrality requirements on all the first stage decisions variables.

In the previous model, one should note that the objective function (**??**) simply includes the recourse function. No first stage cost is considered. Similar to what was is done by **?**, in order to obtain a more efficient implementation of the 0-1 integer L-shaped algorithm (**?**), we opt to use the following alternative formulation:

$$\min_x \tilde{c}x + \tilde{Q}(x) \tag{5.8}$$

where $\tilde{c}_{ij}$, $\forall (v_i, v_j) \in E$, is a lower bound on the expected travel times associated with the edges of the graph. Considering that even though an edge (v_i, v_j) is part of a given *a priori* route, it may still be skipped in the second stage for a particular scenario. Accordingly, $\tilde{Q}(x)$ is defined as:

$$\tilde{Q}(x) = Q(x) - \tilde{c}x$$

which makes the objective function (**??**) equivalent to the original one (**??**). The lower bound $\tilde{c}x$ is used here as a guideline to help the search in the first stage. To obtain $\tilde{c}x$, we use the same technique as proposed by **?**, let:

$$\tilde{c}_{1j} = p_j c_{1j} \qquad j = 2, \ldots, n \tag{5.9}$$

and

$$\tilde{c}_{ij} = p_i p_j c_{ij} + \frac{1}{2}\Big\{ p_i(1-p_j) \min_{k \neq i,j} c_{ik} + p_j(1-p_i) \min_{k \neq i,j} c_{kj} \Big\} \qquad 1 \leq i \tag{5.10}$$

If $i = 1$ the definition in Equation (**??**) is straightforward. If $i > 1$ the bound $\tilde{c}_{ij}$ takes into account the four possibilities of presence or absence of v_i and v_j. If both v_i and v_j are present and (v_i, v_j) appears in the first stage solution, this edge will be visited with probability $p_i p_j$. If only one of the two vertices is present the bound computes the minimum expected travel time of entering or leaving that vertex. In order to avoid double counting, the second term is multiplied by $\frac{1}{2}$. The term associated with both vertices not being present, i.e, $(1 - p_i)(1 - p_j)$, is zero.

Let us now define the recourse function $Q(x)$ which, as previously mentioned, is obtained by summing both the travel times and the total penalties associated with early and late arrivals at customers. Before doing so, it should be noted that penalties not only depend on the subset of customers needing service, they are also determined according to the orientation on which a route is executed (i. e., arrival times at customers are not the same for each orientation). Therefore, for each route that is considered, an orientation is chosen *a priori* such as to minimize penalty costs. In the case of traveling times, route orientation has no influence. Recall that customers are visited according to the imposed sequence, considering that absentees are simply skipped. Consequently, a vehicle ends its route at the depot at the same time regardless of the orientation that is chosen.

Let $Q^{r,\delta}$ denote the expected recourse cost corresponding to route r if the orientation δ is chosen (where δ=1 or $\delta = 2$), which is formulated as follows: $Q^{r,\delta} = Q^r_T + Q^{r,\delta}_P$. Function Q^r_T represents the total average travel time for route r and $Q^{r,\delta}_P$ defines the total average penalties associated with time window deviations for route r if orientation δ is chosen. Then, the recourse cost associated with a given feasible solution can be explicitly formulated as follows:

$$Q(x) = \sum_{r=1}^{m} min\{Q^{r,1}, Q^{r,2}\}$$

Let us now present in detail how functions Q^r_T and $Q^{r,\delta}_P$ are computed. To obtain function $Q^{r,\delta}_P$, we first consider that the penalty for arriving before the established time window is proportional to the associated earliness. As for the penalty associated with late arrivals, it is defined as the tardiness times a given penalty factor β (where β is a real value such that $\beta > 1$). As previously mentioned, the recourse policy that is used considers that upon arrival, regardless of whether on time, early or late, the vehicle starts servicing customers. This assumption allows for an independent treatment of each vertex on a given route.

Given a route r, let us relabel the vertices on the route according to a given orientation δ in the following way: $(v_1 = v_{1_r^\delta}, v_{2_r^\delta}, \ldots, v_{t_r^\delta}, v_{(t+1)_r^\delta} = v_1)$. If one is considering orientation δ for route r, let function $\phi(v_{i_r^\delta})$ define the minimum expected penalty associated with customer $v_{i_r^\delta}$. One should note that the arrival time at a customer is influenced by the presence or absence of its predecessors along the route. This is dictated by the given route r and the chosen orientation δ. Let $\mathcal{A}_{i_r^\delta}$ define the collection of random events where customer i_r^δ requires a visit (i. e., $\mathcal{A}_{i_r^\delta} = \{\omega \in \Omega \mid \xi_{i_r^\delta}(\omega) = 1\}$). We can now define function $Q_P^{r,\delta}$ as follows:

$$Q_P^{r,\delta} = \sum_{i=1}^{t+1} \phi(v_{i_r^\delta})$$

Where for a given route r and orientation δ, the total average penalty cost $Q_P^{r,\delta}$ is the sum of the minimum expected penalty associated with each customer $v_{i_r^\delta}$.

Function $\phi(v_{i_r^\delta})$ is computed by solving an appropriate LP model. To formulate this model, let us define the following parameters:

- w half length of the time window
- β late arrival penalty
- p_ω probability associated with scenario $\omega \in \mathcal{A}_{i_r^\delta}$
- $a_{i_r^\delta}(\omega)$ arrival time at $v_{i_r^\delta}$ considering that scenario $\omega \in \mathcal{A}_{i_r^\delta}$ is observed

It should be noted that the starting time associated with route r is defined such as: $a_{1_r^1}(\omega) = a_{1_r^2}(\omega) = 0$. Furthermore, considering that the total travel time of route r is not dependent on the orientation, then we define arrival times at the end of route r as follows: $a_{(t+1)_r^1}(\omega) = a_{(t+1)_r^2}(\omega) = a_{(t+1)_r}(\omega)$.

We note that the arrival time $a_{i_r^\delta}(\omega)$, is computed according to the subset of customers that are scheduled before $v_{i_r^\delta}$ in route r following orientation δ and given a particular scenario ω. Therefore, considering that vehicles start their service upon arrival at a customer's location, the value $a_{i_r^\delta}(\omega)$ is simply obtained by summing the travel times between customers along route r following orientation δ and taking into account that present customers are defined according to the observed scenario $\omega \in \mathcal{A}_{i_r^\delta}$.

We also define the following variables:

- $t_{i_r}^\delta$ target arrival at $v_{i_r^\delta}$
- $l_{i_r}^\delta(\omega)$ tardiness at $v_{i_r^\delta}$ considering that scenario $\omega \in \mathcal{A}_{i_r^\delta}$ is observed

- $e_{i_r}^{\delta}(\omega)$ earliness at $v_{i_r^{\delta}}$ considering that scenario $\omega \in \mathcal{A}_{i_r^{\delta}}$ is observed

Clearly, targets are dependent on the orientation that is considered. Furthermore, both recourse decisions for tardiness and earliness also vary according to the orientation.

The LP formulation can now be expressed as follows:

$$\phi(v_{i_r^{\delta}}) = \min \sum_{\omega \in \mathcal{A}_{i_r^{\delta}}} p_{\omega}(e_{i_r}^{\delta}(\omega) + \beta l_{i_r}^{\delta}(\omega)) \tag{5.11}$$

subject to

$$[t_{i_r}^{\delta} - w] - a_{i_r^{\delta}}(\omega) \leq e_{i_r}^{\delta}(\omega) \qquad \forall \omega \in \mathcal{A}_{i_r^{\delta}} \tag{5.12}$$

$$a_{i_r^{\delta}}(\omega) - [t_{i_r}^{\delta} + w] \leq l_{i_r}^{\delta}(\omega)) \qquad \forall \omega \in \mathcal{A}_{i_r^{\delta}} \tag{5.13}$$

$$0 \leq e_{i_r}^{\delta}(\omega) \qquad \forall \omega \in \mathcal{A}_{i_r^{\delta}} \tag{5.14}$$

$$0 \leq l_{i_r}^{\delta}(\omega) \qquad \forall \omega \in \mathcal{A}_{i_r^{\delta}} \tag{5.15}$$

$$0 \leq t_{i_r^{\delta}} \tag{5.16}$$

Through the computation of $\phi(v_{i_k^{\delta}})$ in (**??**), one fixes the decision variables $t_{i_r}^{\delta}$, which define the target value for $v_{i_r^{\delta}}$. As previously stated, these targets are presumed to be in the middle of a time window of length $2w$ which is communicated to the customer. The value obtained for $\phi(v_{i_k^{\delta}})$ is a weighted sum of deviations from the established targets. Constraints (**??**) and (**??**) ensure that both earliness and tardiness from the windows are attributed to $e_{i_k}(\omega)$ and $l_{i_k}(\omega)$ respectively. Constraints (**??**) and (**??**) guarantee that for each ω tardiness and earliness cannot have negative values.

We now conclude this section by defining how to compute Q_T^r (i. e., the total average traveling time for route r). Considering the no-wait assumption, i. e., a vehicle immediately starts to service a customer when it arrives at the customer's location, Q_T^r is simply measured by computing the following weighted sum:

$$Q_T^r = \sum_{\omega \in \Omega} p_{\omega} a_{t_r+1}(\omega)$$

which defines the expected arrival time of a vehicle at the depot at the end of route r.

5.3. Solution Methodology

We propose to solve problem (**??**)-(**??**) using the 0-1 integer L-Shaped algorithm developed by **?**, which is used to solve two stage stochastic programming problems with binary first stage and integer recourse. The proposed algorithm is based on the L-shaped method proposed by **?** for continues problems and on Benders decomposition (**?**). The algorithm follows a branch-and-cut solution strategy that solves at each node of the search tree a relaxed model referred to as the current problem (CP). The CP model is initially obtained from (**??**)-(**??**) by relaxing the integrality constraints (**??**), the subtour elimination and route duration constraints (**??**), and by approximating the recourse function $Q(x)$ by a valid lower bound defined by Θ. The general algorithm can now be described as follows (we adopt here the notation used by **?**):

Step 0 Set iteration counter $\nu := 0$ and introduce $\Theta \geq L$ to the CP. Set the value of the best known solution $\bar{z} := +\infty$. The only pendent node is the initial CP.

Step 1 Select a pendent node from the list. If none exists stop.

Step 2 Set $\nu := \nu + 1$ and solve CP. Let (x^ν, Θ^ν) be the optimal solution obtained.

Step 3 Check for any violated constraints of type (**??**). At this stage, valid inequalities or lower bounding functionals may also be generated. If a violated constraint is found add it to the CP and Return to Step 2. Otherwise, if $cx^\nu + \Theta^\nu \geq \bar{z}$, fathom the current node and return to Step 1.

Step 4 If the solution is not integer then branch on a fractional variable. Append corresponding subproblems to the list of pendent nodes and return to Step 1.

Step 5 Compute $Q(x^\nu)$ and set $z^\nu := cx^\nu + Q(x^\nu)$. If $z^\nu < \bar{z}$ then $\bar{z} = z^\nu$.

Step 6 If $\Theta^\nu \geq Q(x^\nu)$, then fathom the current node and return to Step 1. Otherwise add an optimality cut defined as:

$$\sum_{\substack{1<i<j \\ x_{ij}^\nu = 1}} x_{ij} \leq \sum_{1<i<j} x_{ij}^\nu - 1$$

and go to Step 2.

There are two important components within the previous algorithm that have to be clearly defined. The first relates to how the general lower bound L is computed. The strategy that is proposed is developed in subsection **??**. The second component directly concerns how Step 2 of the algorithm is performed. In this step, a search for violated constraints is performed such as to cut, if possible, the considered solution (x^ν, Θ^ν). There are two types of cuts that can be added at this point, either subtour and duration constraints, which enforce feasibility of the routes in the first stage, or lower bounding functionals, which improve the obtained value for Θ. In subsection **??** we define the separation strategies that are proposed for each of the valid inequalities used.

5.3.1 Lower Bound

The general lower bound L is comprised of two components: a general lower bound on the penalties and a general lower bound on the traveling times. The general lower bound on the penalties $Q_{P(L)}$, is presented in subsection ??, it is constructed by defining a valid lower bound for the penalty associated with each customer. The general lower bound on the traveling times $Q_{T(L)}$ is presented in subsection ??, it is established by defining a modified problem which constitutes a lower bound on the traveling times, when compared to the original problem.

5.3.2 Lower Bound on the penalties

The general lower bound on the penalties is constructed in two main steps. First, we present the penalties associated with the first customer scheduled after the depot. Then we bound the penalties associated with the second customer scheduled after the depot. We show that the bound, established for the second scheduled customer, is a valid lower bound for each customer subsequent to the second. Finally, the general lower bound is constructed by combining the bounds.

Given any route r with orientation δ, the penalty associated with the customer scheduled immediately after the depot ($v_{2_r^\delta}$) is zero, i. e., $\phi(v_{i_2^\delta}) = 0$. This can be explained by the fact the arrival time at $v_{2_r^\delta}$ is unaffected by other customers, i. e., $a_{2_r^\delta}(\omega) = c_{1,2_r^\delta}$ for all $\omega \in \mathcal{A}_{2_r^\delta}$. Thus, setting the target $t_{2_r}^\delta$ at $v_{2_r^\delta}$ to $c_{1,2_r^\delta}$ results in $\phi(v_{i_2^\delta}) = 0$.

The arrival time at customer $v_{3_r^\delta}$ is influenced solely by its preceding customer $v_{2_r^\delta}$. Thus, the arrival time at $v_{3_r^\delta}$ can be expressed by conditioning on the existence of $v_{2_r^\delta}$, as follows:

$$a_{3_r^\delta}(\omega) = \begin{cases} c_{1,2_r^\delta} + c_{2_r^\delta,3_r^\delta} & \text{if client } v_{2_r^\delta} \text{ is present in } \omega \in \mathcal{A}_{3_r^\delta} \\ c_{1,3_r^\delta} & \text{if client } v_{2_r^\delta} \text{ is not present in } \omega \in \mathcal{A}_{3_r^\delta} \end{cases}$$

If $w > c_{1,2_r^\delta} + c_{2_r^\delta,3_r^\delta}$ setting the $t_{3_r}^\delta$ to zero would result in $\phi(v_{3_r^\delta}) = 0$. Furthermore, if $c_{1,3_r^\delta} + w > c_{1,2_r^\delta} + c_{2_r^\delta,3_r^\delta} - w$, then $2w > c_{1,2_r^\delta} + c_{2_r^\delta,3_r^\delta} - c_{1,3_r^\delta}$. Thus, setting $t_{3_r}^\delta$ to any value in the interval $[c_{1,3_r^\delta} + w, c_{1,2_r^\delta} + c_{2_r^\delta,3_r^\delta} - w]$, would result in $\phi(v_{3_r^\delta}) = 0$.

For all other cases the following holds:

$$c_{1,3_r^\delta} + w \leq t_{3_r}^\delta \leq c_{1,2_r^\delta} + c_{2_r^\delta,3_r^\delta} - w$$

We rewrite $\phi(v_{3^\delta_r})$ as follows:

$$\phi(v_{3^\delta_r}) = \min_{t^\delta_{3_r}} \left[p_{2^\delta_r}\beta[c_{1,2^\delta_r} + c_{3^\delta_r,2^\delta_r} - t^\delta_{3_r} - w]^+ + q_{2^\delta_r}[t^\delta_{3_r} - w - c_{1,3^\delta_r}]^+\right]$$

As $\beta > 1$, $\phi(v_{3^\delta_r})$ is decreasing with $t^\delta_{3_r}$. Thus, $\phi(v_{3^\delta_r})$ is minimized for $t^\delta_{3_r} = c_{1,2^\delta_r} + c_{3^\delta_r,2^\delta_r} - w$ and $\phi(v_{3^\delta_r}) = q_{2^\delta_r}[t^\delta_{3_r} - w - c_{1,3^\delta_r}]^+$. Next, we establish a general lower bound on $\phi(v_{3^\delta_r})$.

Proposition 5.1 Let $\widetilde{q} = \min_{\forall i \in C}(1 - p_i)$, and let ρ be:

$$\rho = \min_{\forall i \neq j, i,j > 1} c_{1i} + c_{ij} - c_{1j}$$

Define $\tilde{\phi} = \widetilde{q}[\rho - 2w]^+$. Then $\tilde{\phi} \leq \phi(v_{3^\delta_r})$.

Proof:

$$\begin{aligned}
\phi(v_{3^\delta_r}) &= q_{2^\delta_r}[t^\delta_{3_r} - w - c_{1,3^\delta_r}]^+ \\
&= q_{2^\delta_r}[c_{1,2^\delta_r} + c_{3^\delta_r,2^\delta_r} - c_{1,3^\delta_r} - 2w]^+ \\
&\geq \widetilde{q}[c_{1,2^\delta_r} + c_{3^\delta_r,2^\delta_r} - c_{1,3^\delta_r} - 2w]^+ \\
&\geq \widetilde{q}[\rho - 2w]^+ = \tilde{\phi}
\end{aligned}$$

□

Next, we show that $\tilde{\phi}$ is a valid lower bound, on the penalties, for any customer subsequent to the second scheduled customer.

Proposition 5.2 For all routes r and orientations δ, denoted by

$$(v_1 = v_{1^\delta_r}, v_{2^\delta_r}, \ldots, v_{t^\delta_r}, v_{(t+1)^\delta_r} = v_1),$$

there holds the lower bound

$$\phi(v_{k^\delta_r}) \geq \tilde{\phi}, \qquad \forall\, k = 3, \ldots, t+1.$$

Proof: For sake of notation, we drop the indexes r and δ throughout the proof. Let A_k be the collection of random events where customer k requires a visit and let $A_k^c = \Omega \backslash A_k$ be the complement of A_k in Ω. Referring to definition (**??**), let

$$f_k(\omega) = e_k(\omega) + \beta l_k(\omega).$$

We remark that the minimization occurring in the computation of $\phi(v_k)$ in this proof, is always subject to constraints (**??**)-(**??**). For all $k > 3$ we compute

$$\begin{aligned}\phi(v_k) &= \min \sum_{\omega \in A_k} p_\omega f_k(\omega) \\ &= \min \left\{ \sum_{\omega \in A_k \cap A_{k-2}} p_\omega f_k(\omega) \quad + \sum_{\omega \in A_k \cap A_{k-2}^c} p_\omega f_k(\omega) \right\} \\ &\geq \min \sum_{\omega \in A_k \cap A_{k-2}} p_\omega f_k(\omega) \quad + \quad \min \sum_{\omega \in A_k \cap A_{k-2}^c} p_\omega f_k(\omega).\end{aligned}$$

We want to prove that the last line is bounded from below by $\tilde{\phi}$. We proceed in two steps.

Step I. First we prove that for all $k > 3$

$$\Phi_1 = \min \sum_{\omega \in A_k \cap A_{k-2}} p_\omega f_k(\omega) \geq p_{k-2} \tilde{\phi}. \tag{5.17}$$

As in the proof of Proposition **??**,

$$\begin{aligned}\Phi_1 = \min_t p_{k-2} \Big[& q_{k-1} \Big(a_{k-2}(\omega) + c_{(k-2)k} - (t - w) \Big)^- \\ & + \beta p_{k-1} \Big(a_{k-2}(\omega) + c_{(k-2)(k-1)} + c_{(k-1)k} - (t + w) \Big)^+ \Big].\end{aligned}$$

The expression is minimized by $t = a_{k-2}(\omega) + c_{(k-2)(k-1)} + c_{(k-1)k} - w$. Substituting, we get

$$\Phi_1 = p_{k-2} q_{k-1} \big[c_{(k-2)(k-1)} + c_{(k-1)k} - c_{(k-2)k} - 2w \big]^+ \geq p_{k-2} \tilde{q} [\rho - 2w]^+ = p_{k-2} \tilde{\phi}.$$

Step II. Now we prove that

$$\Phi_2 = \min \sum_{\omega \in A_k \cap A_{k-2}^c} p_\omega f_k(\omega) \geq q_{k-2} \tilde{\phi}. \tag{5.18}$$

For all $k > 3$ define the sets

$$B_i = A_k \cap \left(\bigcap_{j=i}^{k-2} A_j^c \right), \qquad i = 2, \ldots, k-2,$$

so that (**??**) can be written as

$$\min \sum_{\omega \in B_{k-2}} p_\omega f_k(\omega) \geq q_{k-2} \tilde{\phi} \qquad \text{for all } k > 3.$$

We proceed by induction *on the index i*. We claim that for all $k > 3$

$$\min \sum_{\omega \in B_i} p_\omega f_k(\omega) \geq \tilde{\phi} \prod_{j=i}^{k-2} q_j \quad \text{for all } i = 2, \dots, k-2.$$

First we prove the induction basis ($i = 2$):

$$\begin{aligned}
\min \sum_{\omega \in B_2} p_\omega f_k(\omega) &\geq \min_t \prod_{j=2}^{k-2} q_j \Big[q_{k-1} \Big(c_{1k} - (t-w) \Big)^- \\
&\qquad + \beta p_{k-1} \Big(c_{1(k-1)} + c_{(k-1)k} - (t+w) \Big)^+ \Big] \\
&\geq \prod_{j=2}^{k-2} q_j [\rho - 2w]^+ q_{k-1} \geq \tilde{\phi} \prod_{j=2}^{k-2} q_j.
\end{aligned}$$

Then, assuming

$$\min \sum_{\omega \in B_i} p_\omega f_k(\omega) \geq \tilde{\phi} \prod_{j=i}^{k-2} q_j,$$

we prove the induction step, i. e., that the claim for i implies the claim for $i+1$. Note that, by definition, $B_i = B_{i+1} \cap A_i^c$.

$$\begin{aligned}
\min \sum_{\omega \in B_{i+1}} p_\omega f_k(\omega) &= \min \left(\sum_{\omega \in B_{i+1} \cap A_i} p_\omega f_k(\omega) + \sum_{\omega \in B_{i+1} \cap A_i^c} p_\omega f_k(\omega) \right) \\
&\geq p_i \prod_{j=i+1}^{k-2} q_j \min_t \Big[q_{k-1} \Big(a_i(\omega) + c_{ik} - (t-w) \Big)^- \\
&\qquad + \beta p_{k-1} \Big(a_i(\omega) + c_{i(k-1)} + c_{(k-1)k} - (t+w) \Big)^+ \Big] \\
&\qquad + \min \sum_{\omega \in B_i} p_\omega f_k(\omega) \\
&\geq p_i \prod_{j=i+1}^{k-2} q_j [\rho - 2w]^+ q_{k-1} + \tilde{\phi} \prod_{j=i}^{k-2} q_j \\
&\geq p_i \tilde{\phi} \prod_{j=i+1}^{k-2} q_j + q_i \tilde{\phi} \prod_{j=i+1}^{k-2} q_j = \tilde{\phi} \prod_{j=i+1}^{k-2} q_j.
\end{aligned}$$

This concludes the proof of Step II. Inequalities (**??**) and (**??**) yield the proof of Proposition **??**. □

Note that the penalty on customer $v_{2^\delta_r}$, $\phi(v_{2^\delta_r})$, is zero. Given n customers and m vehicles, under the problem definition in (??)-(??) each route contains at least a single customer. Thus, for m customers the penalty value will be zero. Proposition ?? shows that for each customer $v_{k^\delta_r}$ such that $k \geq 3$ the following holds: $\phi(v_{k^\delta_r}) \geq \tilde{\phi}$. There are at least $n - m - 1$ customers who are not direct successors of the depot. Thus, the general lower bound for the penalties $Q_{P(L)}$ is presented

$$Q_{P(L)} = (n - m - 1)\tilde{\phi}$$

5.3.3 Lower Bound on the traveling times

The general lower bound on the travel times is constructed by defining a modified problem to the one presented in (??)-(??). Essentially, in the modified problem the distances between every two nodes are presumed equal, for all nodes in the graph. Furthermore, a unique probability of occurrence is assumed for each customer. The fact that distances and probabilities are equivalent for each customer enables a more generalized treatment.

The modified problem is obtained as follows: let us first define the minimum travel time between vertices in the graph $G(V, E)$ as l, where $l = \min\limits_{\forall i \neq j} c_{ij}$. Let us now assume that all traveling times, defined for all $(v_i, v_j) \in E$, are set equal to l. Furthermore, let the probability of occurrence for each customer be fixed to $\widetilde{p}$, where $\widetilde{p} = \min\limits_{\forall i \in C} p_i$; and let the probability of a customer not needing service be set to $\widetilde{q} = 1 - \widetilde{p}$.

For route r, let $Q^r_{T(L)}$ represent the total average travel time in the modified problem. Considering $\widetilde{p}$ as probability of occurrence will result in lowering the expected travel times, when compared to the original problem. This, coupled with the fact that all traveling times are set to l, establishes that $Q^r_{T(L)} < Q^r_T \quad \forall r$.

$Q^r_{T(L)}$ solely depends on the number of customers allocated to the given route r comprised of $(v_1 = v_{1_r}, v_{2_r}, \ldots, v_{t_r}, v_{(t+1)_r} = v_1)$. In this modified version of the problem, the expected route length up to t_r is binomially distributed with t_r and $\widetilde{p}$. Thus, the expected route length can be bounded by $(t_r)\widetilde{p}l$. As such we do not consider one extra distance l to be traversed for each scenario where at least one customer is present. Furthermore, we note that the scenario by which all customers in r are not present has a travel time of zero (not one). Equation (??) provides this expected route length.

$$Q^r_{T(L)} = (t_r)\widetilde{p}l \tag{5.19}$$

From Equation(**??**) it is evident that the expected travel times, for any allocation of $n-1$ customers to m vehicles, yields similar total travel time. Accordingly the lower bound on the total travelling time can be computed as follows:

$$Q_{T(L)} = (n-1)\widetilde{pl}$$

Putting both lower bounds for the penalty and travelling time together we conclude that:

$$L = (n-1)\widetilde{pl} + (n-m-1)\tilde{\phi}$$

5.3.4 Feasibility Cuts and Lower Bounding Functionals

In this section we start by discussing feasibility cuts, for which we adapted a separation procedure from the capacitated VRP literature to fit the route duration constraints. Later we present the lower bounding functional that is a tailored cut for Θ.

Checking violated constraints of type **??** requires the computation of:

$$\ell(S^+) = \sum_{i,j \in S^+ \subset V} c_{ij} x_{ij}$$

The term $\left\lceil \frac{\ell(S^+)}{\lambda} \right\rceil$ denotes the number of vehicles required to visit customers in S corresponding to a specific solution at hand, we recall that $S^+ = S \cup \{1\}$. A more general approach is to bound the required number of vehicles needed to serve S irrespectively of the specific solution. Let $H(S^+)$ be the solution value to the travelling salesman problem (TSP) over S^+. Due to the triangle inequality, $\left\lceil \frac{H(S^+)}{\lambda} \right\rceil$ is a lower bound on the number of vehicles required for servicing S. Thus we use $\left\lceil \frac{H(S^+)}{\lambda} \right\rceil$ for the separation procedure. This is equivalent to the use of rounded capacity constraints RCC in the capacitated VRP. Thus, we call them rounded duration constraints RDC.

For the problem at hand we adapted the connected components heuristic separation procedure proposed by **?**. The heuristic computes connected components $S_1, ..., S_p$. The RDC are checked for each S_i^+ as well as $V \setminus S_i$. Finally, the RDC for the union of those components which are not connected to the depot is checked. Similar heuristics based on connected components are used in **?**. Contrary to RCC, integer solutions are checked by computing the corresponding $\ell(S^+)$, as the $H(S^+)$ value in this case is not guaranteed to find violated constraints.

Next we construct a lower bounding functional LBF on $Q(x)$ via partial routes as introduced by **?**. A partial route h is specified by two ordered sets of vertices $S_h =$

$\{v_1, ..v_{s_h}\}$ and $T_h = \{v_1, ..v_{t_h}\}$ satisfying $S_h \cap T_h = \{v_1\}$. A third set U_h satisfies $S_h \cap U_h = \{v_{s_h}\}$ and $T_h \cap U_h = \{v_{t_h}\}$. We denote $(v_i, v_j) \in S_h$ or T_h if v_i and v_j are consecutive in S_h or T_h. The partial route h is comprised of two chains and some unrestricted vertex set U_h. Essentially these bounds make use of the information from the partial route.

In the context of our problem, given the partial route h based on a specific orientation δ, we bound $Q^{h,\delta}$. Assuming an orientation, the partial route is comprised of three main elements. A sequenced chain, followed by an unrestricted vertex set followed by a sequenced chain. We bound each of these elements, in terms of penalties and travelling time, and later we add these bounds to produce a valid lower bound on $Q^{h,\delta}$. First, we discuss the computation of the penalty competent for each of the three elements in h, later we discuss the bound on the traveling time.

We define P_h^δ as the lower bound for $Q^{h,\delta}$. A partial route h with orientation δ consists of a beginning chain $(v_1 = v_{1_h^\delta}, v_{2_h^\delta}, \ldots, v_{s_h^\delta})$, followed by $\tilde{U}_h = U_h \setminus \{v_{s_h}, v_{t_h}\}$. Finally the partial route ends with the chain $(v_{t_h^\delta}, v_{t-1_h^\delta}, \ldots, v_{1_h^\delta})$. The bound for partial route h is defined as follows:

$$P_h = \min\{P_h^1, P_h^2\}$$

Given the chain $(v_1 = v_{1_h^\delta}, v_{2_h^\delta}, \ldots, v_{s_h^\delta})$, we define the penalties associated with this chain by $Q_P^{S_h,\delta}$. Since this chain is virtually unaffected by U_h the computation of $Q_P^{S_h,\delta}$ is done in an exact as manner as in Section **??**. This is denoted by:

$$Q_P^{S_h,\delta} = \sum_{i=1}^{s_h^\delta} \phi(v_{i_r^\delta}) \tag{5.20}$$

We remark that the minimization occurring in the computation of $\phi(v_{i_r^\delta})$ in this section, is always subject to constraints (**??**)-(**??**).

The penalties associated with $\tilde{U}_h$ are related to having the chain $(v_1 = v_{1_h^\delta}, v_{2_h^\delta}, \ldots, v_{s_h^\delta})$ as a predecessor. The bound on the penalty costs for $\tilde{U}_h$ is denoted by $B_{P(U_h)}^\delta$. This bound is constructed by considering that $v_{s_h^\delta}$ is present. By considering scenarios where $v_{s_h^\delta}$ is present, we account only for a subset of the scenarios, whose associated penalty is at least $\phi(v_{s_h^\delta})$. Under this assumption the penalty of each vertex in $\tilde{U}_h$ is bounded by $p_{s_h^\delta}\phi(v_{s_h^\delta})$. Thus, the $B_{P(U_h)}^\delta$ is computed as follows:

$$B_{P(U_h)}^\delta = |\tilde{U}_h| p_{s_h^\delta} \phi(v_{s_h^\delta}) \tag{5.21}$$

As for bounding the penalties associated with the chain $(v_{t^\delta_h}, v_{t-1^\delta_h}, \ldots, v_{1^\delta_h})$, we denote the number of customer vertices in $(v_{t^\delta_h}, v_{t-1^\delta_h}, \ldots, v_{1^\delta_h})$ by $|T_h|$. Similar to $B^\delta_{P(U_h)}$, the bound $B^\delta_{P(T_h)}$ is constructed by considering that $v_{s^\delta_h}$ is present. Thus, the penalty in $(v_{t^\delta_h}, v_{t-1^\delta_h}, \ldots, v_{1^\delta_h})$ is bounded by $p_{s^\delta_h}\phi(v_{s^\delta_h})$.

$$B^\delta_{P(T_h)} = |T_h| p_{s^\delta_h}\phi(v_{s^\delta_h}) \tag{5.22}$$

The lower bound on the travelling time is constructed by concepts similar to those used for $Q_{T(L)}$ in section **??**. We modify the distances and probabilities from the original problem in order to achieve the bound. In what follows we elaborate on these modifications.

Let the distance between each vertex in $\tilde{U}_h$ and each vertex in S^δ_h be the following:

$$l^\delta_{S_h} = \min_{\forall i \in S^\delta_h j \in \tilde{U}_h} c_{ij}$$

Furthermore, let the probability of occurrence for each customer in $\tilde{U}_h$ be fixed to $\widetilde{p}$, where:

$$\widetilde{p} = \min_{\forall i \in \tilde{U}_h} p_i$$

Let the probablity of a customer in $\tilde{U}_h$ not needing service be set to $\widetilde{q}$, where $\widetilde{q} = 1 - \widetilde{p}$.

Let the distance between each vertex in $\tilde{U}_h$ be the following:

$$l_{\tilde{U}_h} = \min_{\forall i,j \in \tilde{U}_h; i \neq j} c_{ij}$$

Finally, let the distance between each vertex in $\tilde{U}_h$ and each vertex in T^δ_h be the following:

$$l^\delta_{T_h} = \min_{\forall i \in T^\delta_h j \in \tilde{U}_h} c_{ij}$$

Let the aforementioned modifications define the route $\hat{r}^\delta$ which is similar to the partial route h only with sequenced nodes in $\tilde{U}_h$. Clearly, the travelling time in $\hat{r}^\delta$ is a lower bound on the travelling time of h. We define the bound on the total travel time as $B_{T(\hat{r})}$ and it is expressed in Equation **??**.

$$B_{T(\hat{r})} = \sum_{\omega \in \Omega} p_\omega a_{t_{\hat{r}}+1}(\omega) \tag{5.23}$$

Putting together the values from Equations (??) - (??) defines P_h^δ a follows:

$$P_h^\delta = Q_P^{S_h,\delta} + B_{P(U_h)}^\delta + B_{P(T_h)}^\delta + B_{T(\hat{r})}$$

Let $R_h = S_h \cup T_h \cup U_h$, $x = (x_{ij})$ and

$$W_h(x) = \sum_{(v_i,v_j)\in S_h} x_{ij} + \sum_{(v_i,v_j)\in T_h} x_{ij} + \sum_{(v_i,v_j)\in U_h} x_{ij} - |R_h| + 1$$

Given r partial routes, let P_{r+1} be a lower bound for $Q(x)$ for $m-r$ routes involving customers $V \setminus \cup_{h=1}^r R_h$. Let $P = \sum_{h=1}^{r+1} P_h$ and L be the general lower bound. Similar to ? constraint (??) is a valid inequality for SConVRP.

$$\Theta \leq L + (P + L)\left(\sum_{h=1}^{r} W_h(x) - r + 1\right) \tag{5.24}$$

We note that in our implementation P_{r+1} is computed in a similar fashion as L, while accounting only for $V \setminus \cup_{h=1}^r R_h$ customers.

We conclude this section by explaining how these valid inequalities are generated. In step 3 of the 0-1 integer L-Shaped algorithm presented, we start by searching for violated subtour elimination and route duration constraints. This is done by adapting the CVRPSEP (?) package while using a dynamic programming algorithm for solving the TSP. These cuts are added as long as violated constrains are found. When this heuristic fails, the separation continues by finding violated inequalities of type (??).

5.4. Experiments

To test SConVRP a number of test problems were generated. The problems were generated based on the same principles as in ?. Namely, vertices were generated in $[0, 100]^2$ following a uniform distribution. The customer probability of occurence was randomly set to one out of $[0.6, 0.75, 0.9]$. The branching procedure was based on the package proposed by ?. The experiments were performed on a 2.4 GHz AMD Operton 64 bit processor.

We conducted experiments on 16 data sets. The number of nodes N was set either to 15 or 20. λ (the route duration restriction) was set to 180. The penalty parameter β was set to 2 and w was set to 20. For the sets with 15 nodes the number of vehicles m was set to 3, while for the sets with 20 nodes the number of vehicles m was set to 4. The CPU limit was set to 1800 seconds. The stopping optimality gap for the algorithm was 1%.

Set	N	Initial integer	Gap with Initial integer	$\tilde{c}x$ / Best integer	Final Gap	Run time (sec)
1	15	342.2	8.4%	110%	<1%	74
2	15	453.3	11.6%	113%	<1%	189
3	15	383.1	7.1%	123%	<1%	15
4	15	485.0	14.7%	120%	<1%	173
5	15	461.5	5.9%	138%	<1%	8
6	15	356.5	14.2%	114%	<1%	102
7	15	323.3	9.8%	136%	<1%	90
8	15	442.8	11%	104%	<1%	44
9	20	497.0	4%	134%	<1%	185
10	20	591.8	4%	126%	<1%	91
11	20	545.8	9%	119%	<1%	1670
12	20	409.5	6%	134%	<1%	1914
13	20	505.2	14%	112%	3.46%	1800
14	20	557.4	3%	130%	<1%	222
15	20	516.7	16%	127%	3.58 %	1800
16	20	466.4	6%	133%	2.97 %	1800

Table 5.1 SConVRP results

Table **??** summarizes the results. In 13 out of the 16 cases the algorithm reached an optimality gap of less than 1%. A substantial difference in running times is observed between sets with 15 nodes as opposed to sets with 20 nodes. Sets with 15 nodes ran at most 189 seconds, while sets with 20 nodes required more computation time. Nonetheless, all sets achieved an optimality gap of less than 3.6%. For the problem at hand these computation times are affordable, since the *a priori* plan is expected to serve the daily carrier operations for a long period.

The fifth column reports the gap between the lower bound used in the first stage ($\tilde{c}x$) and the best integer solution. We note that the gaps are relatively high, this is correlated with the run times. Thus, improving the first stage bound will improve the performance of the algorithm.

We note that assuming fully stochastic customers, as is done in this paper, results in extremely difficult problems. It is reasonable that in a realistic setting only a subset of customers are stochastic, meaning that some customers request service on a daily basis while others do not. Thus, if only a subset of customers were presumed stochastic much larger instances may be solved. Furthermore, our solution approach considers all scenarios with respect to the presence and absence of customers. Larger sets might be solved by sampling approaches.

To our knowledge **?** were the only ones who developed separation procedures for route duration constraints. We have adapted a single separation for these constraints, developing additional separation procedures will yield better results. Finally, improving the bounds proposed in sections Sections **??** and **??** will enhance the results.

5.5. Conclusions

We have proposed a stochastic programming formulation for the SConVRP. The consistent VRP was recently identified as a focal problem in the parcel delivery industry. Two main dimensions of consistency are identified in the literature, driver and temporal consistency. Our formulation handles both these dimensions. In our solution approach, driver consistency is preserved while deviations from temporal consistency are penalized.

We developed an exact solution approach for SConVRP. We adopted the 0-1 integer L-shaped algorithm as a solution method. This included defining a general lower bound as well as LBFs. The algorithm was tested on a number of small sized instance (15 to 20 customers). The results indicate that our solution method was successful in solving most test instances. These problems can be used as benchmarks for future research.

Chapter 6

Conclusions

The VRP is central to the operations of most carrier companies. Increasing demand for distribution services, coupled with the need to maintain high customer service, pose significant challenges for these companies. One of the main challenges is to balance operational costs and customer service, while the former is rather explicit to assess, the latter is more involved.

The uncertainties encountered by carrier contribute to increased complexity in making routing decisions. When necessary, uncertainties should be taken into account to ensure sound decisions. These uncertainties may result in substantially diverse routing policies. To address uncertainty, this thesis adopts an *a priori* approach. In particular, solutions are established in the beginning of the planning period and are not re-optimized in accordance with random events. The impact of a random occurrence is estimated in terms of the additional cost it imposes on the solution.

Transportation of goods is key to society. However, there are many undesirable side effects to consider. These include increasing emissions, traffic congestion, and noise levels. Moreover, the transport sector is subject to exogenous constraints, e. g., speed limits. Such constraints, might come in the form of governmental, regional, or municipal regulations. Carrier companies need to obey to current regulations and prepare themselves for future ones to come. On the other hand, carrier companies may choose to consider these aspect in their operations planning as part of a social responsibility agenda. In certain situations, considering such societal aspects in the operations planning may lead to profitability. In this thesis we analyze the impact of incorporating CO_2 emissions in a number of VRP settings, we demonstrate situations where reducing CO_2 does not contradict with cost minimization.

In this thesis, we studied four different variants of the VRP. We focused on two distinct dimensions, time and timing. Time-dependent travel time regularly appears in road networks in the form of congestion. Thus, a more realistic estimate of travel time calls for the inclusion of time-dependent travel times. This enables a more enhanced estimation of travel times and of their associated costs.

The inclusion of time-dependent travel times imposes a number of additional elements to the routing problem. The question of which vehicle speed minimizes CO_2 emissions is trivial in a time independent setting. However, considering time-dependent travel times, we showed that the optimal speed could be higher, when compared to its VRP counterpart, in order to avoid congestion. We considered disruptions in customer service times. Again, this question is trivial in a time-independent setting as the optimal solution will not cease to be so if disrupted. However, in a time-dependent environment disruption might substantially influence the performance of the solution. This influence is manifested by the fact that if a disruption occurs in congestion periods, it may cause a vehicle to eventually depart from a customer location in a non-congested period. Thus, the impact of a disruption on travel may be less than the disruption value itself.

In the context of timing in VRP, this thesis introduces SITW (Self-Imposed Time Windows). This is observed in the operations of a number of companies. SITW relate to situation where a time window is communicated to the customer by the carrier. Violations of these windows are penalized. Determining SITW, i. e., the timing of arrivals at customers, is an interesting question in a stochastic environment. The importance of SITW lies in TW being part of the routing decisions. As such, the VRP-SITW can be viewed as an extension of the VRP to include a customer service perspective. Furthermore, the VRP-SITW can be viewed as a relaxed version of VRPTW, where in the latter time windows are exogenous constraints on the routing problem.

First, we introduce the VRP-SITW where disruptions in travel times are considered. We propose coping with such disruptions by inserting buffers between arrival times. For the presented experimental settings, we showed that from a total travel time perspective the VRP-SITW does not entail substantial increases when compared to the VRP.

We integrate SITW to a newly introduced customer service perspective, the Consistent VRP. The definition of consistency stems from the needs expressed by UPS. According to this definition, consistency is having the same driver visiting the same customers at roughly the same time on each day that these customers require service. SITW were used to maintain this temporal consistency. We consider stochastic customers, i. e., customers which may, or may not, be present at a given day. In this setting, SITW played a crucial role in the long term consistent customer service aspect.

The remainder of this chapter is organized as follows, in section **??** we discuss the main conclusions from Chapters 2-5. In Section **??** we summarize directions for future research.

6.1. Discussion

Incorporating time and timing in realistic aspects of the VRP is the overall main contribution of this thesis. Within the broad title of time and timing we have identified a number of realistic problems and embedded them in a VRP framework. We presented models for each of these problems and developed both heuristic and exact solution procedures. We analyzed the resulting solutions, compared them to standard solutions, and derived insights where possible.

In Chapters 3-5 we considered three types of uncertainties. In Chapter 3 stochastic service times was addressed. In Chapter 4 a form of stochasticity in travel times was modeled. Finally, in Chapter 5 stochastic customers were considered.

Chapter 2 studied the TDVRP in conjuncture with CO_2 emissions. We presented a model that considers travel time, fuel, and CO_2 emission costs. Specifically, we proposed a framework for modeling CO_2 emissions in a time-dependent VRP context (E-TDVRP). Since there is a clear correlation between vehicle speed and the amount of CO_2 emissions, the vehicle speed limit was considered as a decision variable. The CO_2 emissions were derived as a function of speed, the function used had a unique speed at which the level of CO_2 emissions per kilometer is minimized. Different speed limits affected both travel time and CO_2 emissions. Lowering vehicle speed led to increasing total travel time. However, the impact of lowering vehicle speed on CO_2 emissions was less straightforward, as it implied spending more time in congestion, which resulted in high emissions. Thus, all performed experiments indicated that from a CO_2 emissions standpoint it is better to set vehicle speed higher than the speed which minimizes emissions per kilometer.

CO_2 emissions are weighed against the traditional travel time cost. Considering a vehicle speed limit of 90 km/hr in the TDVRP, our experiments showed that an average reduction of 11.4% in CO_2 emissions can be achieved. However, this reduction increased travel times by an average of 17.7%. Considering the speed limit of 80 km/hr, the trade-off was less substantial, since an average increase of 6.6% in travel achieved a reduction of 3.2%in CO_2 emissions.

The other important contribution lies in incorporating fuel costs in the optimization. As fuel costs are correlated with CO_2 emissions, Chapter 2 shows that even in today's cost structure limiting vehicle speeds is beneficial. For a number of situations, we show that limiting vehicle speeds is desired from a total cost perspective. This namely stems from the trade-off between fuel and travel time costs. In all presented experiments, the resulting speed limit was less than 90 km/hr.

Considering time-dependent travel times, Chapter 3 dealt with the perturbed TDVRP (P-TDVRP). Delays at customer locations were modeled as disruptions or perturbations. The optimization used a standard tabu search algorithm and altered the evaluation of solutions to include the expected travel time under disruption.

Comparing the trade-off between P-TDVRP and TDVRP solutions, we showed that in more than half of the 108 experimental instances the trade-off was positive. This implies that the benefits of using the P-TDVRP outweighed the additional travel time required by these solutions, in comparison to the TDVRP solutions.

In Chapter 3, we elaborated on the absorption effect (AE). It is calculated as the difference between the perceived perturbed travel time and the actual travel time. We showed that positive AE happens in situations where service time disruptions occur in congestion periods. In this case, the actual departure time from a customer is shifted to a non-congested period. In 96 out of the 108 experiments the P-TDVRP solutions exhibited positive AE values. Furthermore, we identified situations that are more prone to absorb disruptions.

In Chapter 4, self-imposed time windows were introduced to the VRP, resulting in the VRP-SITW. The problem corresponds to situations where carrier companies quote their arrival times to their customers. The VRP-SITW broadens the scope of the VRP by incorporating customer service considerations. The considered VRP-SITW was modeled in a stochastic environment, where the carrier company seeks to uphold its quoted time windows as much as possible.

The objective of the VRP-SITW was minimizing the travel and delay costs corresponding to late arrivals at customers. The solution approach, in this chapter, embedded a LP model within a tabu search procedure. Results from the VRP-SITW were compared with results from the standard VRP. These results indicated that the VRP-SITW requires on average a 3% increase in distance. Comparing VRP-SITW with VRPTW solutions showed mixed trends. In the random sets VRP-SITW resulted in substantial decreases both in the number of vehicles and in travel cost. However, the results of the clustered sets did not differ much neither in terms of the number of vehicles, nor of travel cost.

Chapter 5 dealt with the consistent vehicle problem. It focused on providing consistent customer service, since this was identified as crucial by parcel companies. Driver consistency is preserved by enforcing that the same driver visits customers when they request service. Another aspect of consistent service was identified as temporal consistency, which aims to provide customer service in more or less the same time when this service is required. This was reflected by the use of SITW as a means to guarantee temporal consistency.

The existing models for the consistent VRP assume periodic knowledge with respect to customer demand, i. e., given the demand for a certain period, consistent routes are designed. Chapter 5 introduced a stochastic programming formulation for the consistent VRP with stochastic customers. This formulation relaxes the assumption of known demand. An exact solution method is proposed by adapting the 0-1 integer L-shaped algorithm. The majority of test instances reached an optimality gap of less than 1%. Thus, the results from this chapter can be used as benchmarks for future research in the field of SConVRP.

6.2. Future Research

As usual, this thesis raises more questions than answers. We discuss general directions for future research regarding the underlying concepts presented in the thesis.

The models proposed in Chapters 3-5 establish *a priori* solutions for stochastic environments. A major extension could be done by transforming these models in a manner that accommodates dynamic decision making. In Chapter 3, this would require re-optimizing upon realizing disruptions. In Chapter 4, this can be in the form of quoting time windows for a dynamic order arrival setting. In Chapter 5, maintaining consistency while customer orders arrive dynamically throughout the day would be a challenging extension.

Building upon the presented research, further research on a more tactical level can be conducted. Incorporating CO_2 considerations upon fleet acquisition decisions is a relevant research question. The decisions concerning fleet mix and size are relevant questions in the context of SITW. Generally, there is a trade-off between the number of used vehicles and the ability to uphold SITW. Thus, modeling the impact of additional or different vehicles in terms of SITW remains an open question. Morever, as the proposed models depict realistic settings, they may be embedded as sub-problems in the context of rich vehicle routing problems.

In all presented models the evaluation of the target function was rather involved, compared to its VRP counterpart, i. e., total distance. The difference is due to encompassing time and timing aspects. Thus, further investigation of solution structures that yield better outcomes in the proposed models may enhance the presented solution methods. Such enhancement may lead to encompassing additional stochastic features.

Combining SITW with time-dependent travel times is a natural extension to this thesis. In the context of SITW, further research can be directed towards allowing, at a cost, for flexibility in setting time windows. Another extension might focus on a setting where only a subset of customers have fixed time windows.

Bibliography

Aarts, E. H. L., J.K. Lenstra. 1997. *Local Search in Combinatorial Optimization*. Wiley, Chichester, UK.

Applegate, D.L., R.E. Bixby, V. Chvátal, W.J. Cook. 2007. *Traveling salesman problem, A computational study*. PrincetonUniversity Press, Princeton, NJ.

Augerat, P., J.M. Belenguer, E. Benavent, A. Corber, D. Naddef. 1998. Separating capacity constraints in the CVRP using tabu search. *European Journal of Operational Research* **106** 546–557.

Ballestín, F., R. Leus. 2008. Meta-heuristics for stable scheduling on a single machine. *Computers & Operations Research* **35** 2175–2192.

Baumgartner, M., J. Lonardia, Oliver Kruscha. 2008. Improving computerized routing and scheduling and vehicle telematics: A qualitative survey. *Transportation Research Part D* **13** 377–382.

Ben-Tal, A., A. Nemirovski. 1998. Robust convex optimization. *Mathematics of Operations Research* **23**(4) 769–805.

Ben-Tal, A., A. Nemirovski. 1999. Robust solutions to uncertain programs. *Operations research letters* **25** 1–13.

Benders, J. F. 1962. Partitioning procedures for solving mixed-variables programming problems. *Numerische Mathematik* **4** 238–252.

Bertsimas, D.J., D. Simchi-Levi. 1996. A new generation of vehicle routing research: Robust algorithms, addressing uncertainty. *Operations Research* **44**(2) 286–304.

Bräysy, O., M. Gendreau. 2005a. Vehicle routing problem with time windows, part I: Route construction and local search algorithms. *Transportation Science* **39**(1) 104–118.

Bräysy, O., M. Gendreau. 2005b. Vehicle routing problem with time windows, part II: Metaheuristics. *Transportation Science* **39**(1) 119–139.

Cairns, S. 1999. The home delivery of shopping: the environmental consequences, ESRC Transport Studies Unit, University of London. TSU Working Paper.

Clarke, G., J.W. Wright. 1964. Scheduling of vehicles from a central depot to a number of delivery points. *Operations Research* **12**(4) 568–581.

Corbett, C. J., P. R. Kleindorfer. 2001a. Environmental management and operations: Introduction to part 2 (integrating operations and environmental management systems). *Production and Operations Management* **10** 225–227.

Corbett, C. J., P. R. Kleindorfer. 2001b. Environmental management and operations management: Introduction to part 1 (manufacturing and eco-logistics). *Production and Operations Management* **10** 107–111.

Cordeau, J.-F., G. Laporte. 2005. Tabu search heuristics for the vehicle routing problem. *Metaheuristic Optimization via Memory and Evolution*, vol. 30. Operations Research/Computer Science Interfaces Series, 145–163.

Cordeau, J.-F., G. Laporte, A. Mercier. 2003. A unified tabu search heuristic for vehicle routing problems with time windows. *The Journal of the Operational Research Society* **52**(8) 928–936.

Coyle, J.J., E.J. Bardi, J.J. Langley Jr. 1996. *The Management of Business Logistics*. West Publishing.

Daniels, R., J. Carrillo. 1997. β-robust scheduling for single-machine systems with uncertain processing times. *IIE Transactions* **29** 977–985.

Daniels, R., P. Kouvelis. 1995. Robust scheduling to hedge against processing time uncertainty in single-stage production. *Management Science* **41** 363–376.

Dantzig, G.B., J.H. Ramser. 1959. The truck dispatching problem. *Management science* **2** 80–91.

Ericsson, E., H. Larsson, K. Brundell-Freij. 2006. Optimizing route choice for lowest fuel consumption - potential effects of a new driver support tool. *Transportation Research Part C* **14** 369–383.

European Commission. 1999. Methodology for calculating transport emissions and energy consumption. URL `http://www.inrets.fr/ur/lte/cost319/M22.pdf`. Last accessed October 2010.

European Environment Agency . 2007. Transport and environment: on the way to a new common transport policy. *EEA Report* **1**.

Finke, D.A., D.J. Medeiros, M.T.Traband. 2007. Multiple machine JIT scheduling: a tabu search approach. *International Journal of Production Research* **45**(21) 4899–

4915.

Fleischmann, B., M. Gietz, S. Gnutzmann. 2004. Time-varying travel times in vehicle routing. *Transportation Science* **38**(2) 160–173.

Flisberg, P., B. Lidéna, M. Rönnqvist. 2009. A hybrid method based on linear programming and tabu search for routing of logging trucks. *Computers & Operations Research* **36**(4) 1122–1144.

Garcia, B.L., J.Y. Potvin, J.M. Rousseau. 1994. A parallel implementation of the tabu search heuristic for vehicle routing problems with time window constraints. *Computers & Operations Research* **21**(9) 1025–1033.

Gendreau, M., A. Hertz, G. Laporte. 1994. A tabu search heuristic for the vehicle routing problem. *Management Science* **40**(10) 1276–1290.

Gendreau, M., G. Laporte, J.-Y. Potvin. 2002. Metaheuristics for the VRP. P. Toth, D. Vigo, eds., *The vehicle routing problem*. SIAM Monographs on discrete mathematics and applications, Philadelphia, 129–154.

Gendreau, M., G. Laporte, R. Séguin. 1995a. An exact algorithm for the vehicle routing problem with stochastic demands and customers. *Transportation Science* **29**(2) 143–155.

Gendreau, M., G. Laporte, R. Séguin. 1996a. Stochastic vehicle routing. *European Journal of Operational Research* **88**(1) 3–12.

Gendreau, M., G. Laporte, R. Séguin. 1996b. A tabu search heuristic for the vehicle routing problem with stochastic demands and customers. *Operations Research* **44**(3) 469–477.

Gendreau, M, G. Laporte, M. M. Solomon. 1995b. Single-vehicle routing and scheduling to minimize the number of delays. *Transportation Science* **29**(1) 56–62.

Gendron, B., T.G. Crainic, A. Frangioni, F. Guertin. 2005. Oobb: Object-oriented tools for parallel branch-and-bound. *In PAREAO, Mont-Tremblant, Canada* **January** 16–21.

Ghiani, G., G. Laporte, Musmanno. 2004. *Introduction to Logistics Systems Planning and Control*. Wiley & Sons Ltd.

Gillett, B.E., L.R. Miller. 1974. A heuristic algorithm for the vehicle-dispatch problem. *Operations Research* **22**(2) 340–349.

Glover, F. 1989. Tabu search part I. *ORSA Journal of Computing* **1**(3) 190–206.

Glover, F. 1990. Tabu search, part II. *ORSA Journal of Computing* **2**(1) 4–32.

Groër, C., B. Golden, E. Wasil. 2009. The consistent vehicle routing problem. *Manufacturing Service Operations Management* **11**(4) 630–643.

Hadjiconstantinou, E., D. Roberts. 2002. Routing under uncertainty: An application in the scheduling of field service engineers. P. Toth, D. Vigo, eds., *The vehicle routing problem*. SIAM Monographs on discrete mathematics and applications, Philadelphia, 331–350.

Hashimoto, H., M. Yagiura, T. Ibaraki. 2008. An iterated local search algorithm for the time-dependent vehicle routing problem with time windows. *Discrete Optimization* **5** 434–456.

Herroelen, W., R. Leus. 2004. The construction of stable project baseline schedules. *European Journal of Operational Research* **156** 550–565.

Hertz, A., G. Laporte, M. Mittaz. 2000. A tabu search heuristic for the capacitated arc routing problem. *Operations Research* **48**(1) 129–135.

Hill, A.V., W.C. Benton. 1992. Modeling intra-city time-dependent travel speeds for vehicle scheduling problems. *European Journal of Operational Research* **43**(4) 343–351.

Hjorring, C., J. Holt. 1999. New optimality cuts for a single-vehicle stochastic routing problem. *Annals of Operations Research* **86** 569–584.

Hopp, W.J., M.L. Spearman. 1996. *Factory Physics, Foundations for Manufacturing Management*. Mc-Graw Hill.

Ichoua, S., M. Gendreau, J-Y. Potvin. 2003. Vehicle dispatching with time-dependent travel times. *European Journal of Operational Research* **144** 379–396.

Jabali, O., T. Van Woensel, A.G. de Kok, C. Lecluyse, H. Peremans. 2009. Time-dependent vehicle routing subject to time delay perturbations. *IIE Transactions* **41**(12) 1049–1066.

Jung, S., A. Haghani. 2001. A genetic algorithm for the time dependent vehicle routing problem. *Transportation research record. Journal of Transportation research board* **1771** 161–171.

Kirby, H.R., B. Hutton, R.W. McQuaid, R. Raeside, X. Zhang. 2000. Modelling the effects of transport policy levers on fuel efficiency and national fuel consumption. *Transportation Research Part D* **5** 265–282.

Kleindorfer, P. R., K. Singhal, L. N. Van Wassenhove. 2001. Sustainable operations management. *Production and Operations Management* **14** 482–492.

Kouvelis, P., R. Daniels, G. Vairaktarakis. 2000. Robust scheduling of a two-machine flow shop with uncertain processing times. *IIE Transactions* **32** 421–432.

Kouvelis, P., G. Yu. 1997. *Robust Discrete Optimization and its Applications*. Kluwer Academic Publishers.

Kramer, K.J., H.C. Moll, S. Nonhebel, H C. Wilting. 1999. Greenhouse gas emissions related to dutch food consumption. *Energy Policy* **27** 203–216.

Laporte, G. 1992. The vehicle routing problem: An overview of exact and approximate algorithms. *European Journal of Operational Research* **59**(3) 345–358.

Laporte, G. 2007. What you should know about the vehicle routing problem. *Naval Research Logistics* **54** 811–819.

Laporte, G., F. Louveaux, H. Mercure. 1992. The vehicle routing problem with stochastic travel times. *Transportation Science* **26**(3) 161–170.

Laporte, G., F. Louveaux, L. van Hamme. 2002. An intger l- shaped algorithm for the capacitated vehicle routing problem with stochastic demands. *Operations Research* **50**(3) 415–423.

Laporte, G., F. V. Louveaux. 1993. The integer l-shaped method for stochastic integer programs with complete recourse. *Operations Research Letters* **13** 133–142.

Laporte, G., Y. Nobert, M. Desrochers. 1985. Optimal routing under capacity and distance restrictions. *Operations Research* **33**(5) 1050–1073.

Laporte, G., F. Semet. 2002. Classical heuristics for the capacitated VRP. P. Toth, D. Vigo, eds., *The vehicle routing problem*. SIAM Monographs on discrete mathematics and applications, Philadelphia, 109–128.

Lecluyse, C., T. Van Woensel, H. Peremans. 2009. Vehicle routing with stochastic time-dependent travel times. *4OR, A quarterly Journal of Operations Research* **7**(4) 363–377.

Léonardi, J., M. Baumgartner. 2004. CO_2 efficiency in road freight transportation: Status quo, measures and potential. *Transportation Research Part D* **9** 451–464.

Leus, R., W. Herroelen. 2005. The complexity of machine scheduling for stability with a single disrupted job. *Operations Research Letters* **33** 151–156.

Leus, R., W. Herroelen. 2007. Scheduling for stability in single-machine production systems. *Journal of Scheduling* **10** 223–235.

Li, X., P. Tian, S.C.H Leung. 2010. Vehicle routing problems with time windows and stochastic travel and service times: Models and algorithm. *International Journal* *of Production Economics* **125** 137–145.

Lysgaard, J., A. N. Letchford, R. W. Eglese. 2004. A new branch-and-cut algorithm for the capacitated vehicle routing problem. *Mathematical Programming* **100**(2) 423–445.

Malandraki, C., M.S. Daskin. 1992. Time dependent vehicle routing problems: Formulations, properties and heuristic algorithms. *Transportation science* **26**(3) 185–200.

Malandraki, C., R.B. Dial. 1996. A restricted dynamic programming heuristic algorithm for the time dependent traveling salesman problem. *European Journal of Operational Research* **90** 45–55.

Mitrović-Minć, S., G. Laporte. 2004. Waiting strategies for the dynamic pickup and delivery problem with time windows. *Transportation Research Part B* **38** 635–655.

Montemanni, R., J.Bart, M. Mastrolilli, L.M. Gambardella. 2007. The robust traveling salesman problem with interval data. *Transportation Science* **41**(3) 366–381.

Nannicini, G., P. Baptiste, G. Barbier, D. Krob, L. Liberti. 2008. Fast paths in large-scale dynamic road networks. *Computational Optimization and Applications* **45**(1) 143–158.

Or, I. 1976. *Traveling salesman-type combinatorial poroblems and their relation to the logistics of regional blood banking*. Ph.D. Thesis, Northwestern University, Evanston, Illinois.

Osman, I. H. 1993. Metastrategy simulated annealing and tabu search algorithms for the vehicle routing problem. *Annals of Operations Research* **41** 421–451.

Palmer, A. 2008. The development of an integrated routing and carbon dioxide emissions model for goods vehicles. Ph.D. thesis, Cranfield University, United Kingdom. URL http://dspace.lib.cranfield.ac.uk/handle/1826/2547.

Point Carbon. 2010. Trading analytics. URL `http://www.pointcarbon.com/trading`. Last accessed October 2010.

Potvin, J.Y., J.M. Rousseau. 1995. An exchange heuristic for routing problems with time windows. *Journal of the Operational Research Society* **46**(12) 1433–1446.

Puchinger, J., G.R. Raidl. 2005. Combining metaheuristics and exact algorithms in combinatorial optimization: A survey and classification. J. Mira, J.R. Álvarez, eds., *Artificial Intelligence and Knowledge Engineering Applications: A Bioinspired Approach. Lecture Notes in Computer Science*, vol. 3562. Springer, 41–53.

Ralphs, T. 2010. Branch cut and price resource web. URL http://www.branchandcut.org. Last accessed October 2010.

Ralphs, T. K., L. Kopman, W. R. Pulleyblank, L. E. Trotter. 2003. On the capacitated vehicle routing problem. *Mathematical Programming* **94** 343–359.

Sansò, B., F. Soumis, M. Gendreau. 1991. On the evaluation of telecommunications network reliability using routing models. *IEEE Transactions on Communications* **39**(10) 1494–1501.

Sbihi, A., R. W. Eglese. 2007. Combinatorial optimization and green logistics. *4OR: A Quarterly Journal of Operations Research* **5** 99–119.

Schrank, D., T. Lomax. 2009. *Urban Mobility Report*. Texas Transportation Institute, The Texas A&M University System.

Singha, A., S. Gangopadhyaya, P.K. Nandaa, S. Bhattacharyab, C. Sharmab, C. Bhana. 2004. Trends of greenhouse gas emissions from the road transport sector in India. *Science of the Total Environment* **390** 124–131.

Skabardonis, A., P. P. Varaiya, K F. Petty. 2003. Measuring recurrent and non recurrent traffic congestion. *Transportation Research Record* **1856** 118–124.

Slyke, AR. M. Van, R. Wets. 1969. L-shaped linear programs with applications to optimal control and stochastic programming. *SIAM Journal on Applied Mathematics* **17** 638–663.

Smilowitz, K., M. Nowak, T. Jiang. 2009. Workforce management in periodic delivery operations. Department of Industrial Engineering and Management Sciences, Working Paper 09-004.

Solomon, M. M. 2010. VRPTW Benchmark Problems. URL http://web.cba.neu.edu/∼msolomon/problems.htm. Last accessed April 2010.

Solomon, M.M. 1987. Algorithms for the vehicle routing and scheduling problems with time window constraints. *Operations Research* **35**(2) 254–265.

Sörensen, K. 2004. Finding robust solutions using local serch. *Journal of Mathematical Modelling and Algorithms* **3**(1) 89–103.

Springer, U. 2004. The market for tradable GHG permits under the Kyoto Protocol: a survey of model studies. *Energy Economics* **25**(5) 527–551.

Sugawara, S., D.A. Niemeier. 2002. How much can vehicle emissions be reduced? exploratory analysis of an upper boundary using an emissions-optimized trip assignment. *Transportation Research Record* **1815** 29 – 37.

Sungur, I., F. Ordóñez, M. Dessouky. 2008. A robust optimization approach for the capacitated vehicle routing problem with demand uncertainty. *IIE Transactions* **40** 509–523.

Sungur, I., Y. Ren, F. Ordez, M. Dessouky, H. Zhong. 2010. A model and algorithm for the courier delivery problem with uncertainty. *Transportation Science* **44** 193–205.

Taillard, E., P. Badeau, M. Gendreau, F. Guertin, J.Y. Potvin. 1997. A tabu search heuristic for the vehicle routing problem with soft time windows. *Computers & Operations Research* **31**(2) 170–186.

Taillard, E.D. 1993. Parallel iterative serach methods for vehicle routing problem. *Networks* **23** 661–673.

Taniguchi, E., R.G. Thompson, T. Yamada, R. Van Duin. 2001. *City Logistics: Network Modelling and Intelligent Transport Systems*. Pergamon.

Toth, P., D. Vigo. 2002. *The vehicle routing problem*. SIAM Monographs on discrete mathematics and applications, Philadelphia.

Tsutsui, S., A. Ghosh. 1997. Genetic algorithms with a robust solution searching scheme. *IEEE Transactions on Evolutionary Computation* **1**(3) 201–208.

Van Woensel, T. 2003. *Modelling Uninterrupted Traffic Flows, a Queueing Approach*. Ph.D. Dissertation, University of Antwerp, Belgium.

Van Woensel, T., R. Creten, N. Vandaele. 2001. Managing the environmental externalities of traffic logistics: The issue of emissions. *Production and Operations Management* **10**(2) 207–223.

Van Woensel, T., L. Kerbache, H. Peremans, N. Vandaele. 2008. Vehicle routing with dynamic travel times: A queueing approach. *European Journal of Operational Research* **186**(3) 990–1007.

Van Woensel, T., N. Vandaele. 2006. Empirical validation of a queueing approach to uninterrupted traffic flows. *4OR, A quarterly Journal of Operations Research* **4**(1) 59–72.

Vanek, F.M., J.B. Campbell. 1999. Uk road freight energy use by product: trends and analysis from 1985 to 1995. *Transport Policy* **6** 237–246.

Xiang, S., C. Chu, H. Chen. 2008. The study of a dynamic dial-a-ride problem under time-dependent and stochastic environments. *European Journal of Operational Research* **185**(2) 534–551.

Yang, C., D. McCollum, R. McCarthy, W. Leighty. 2008. Meeting an 80 % reduction in greenhouse gas emissions from transportation by 2050: A case study in california. *Transportation Research Part D* **14** 147–156.

Zhong, H., R.W. Hall, M. Dessouky. 2007. Territory planning and vehicle dispatching with driver learning. *Transportation Science* **41** 74–89.

Summary

Time and Timing in Vehicle Routing Problems

The distribution of goods to a set of geographically dispersed customers is a common problem faced by carrier companies, well-known as the Vehicle Routing Problem (VRP). The VRP consists of finding an optimal set of routes that minimizes total travel times for a given number of vehicles with a fixed capacity. Given the demand of each customer and a depot, the optimal set of routes should adhere to the following conditions:

- Each customer is visited exactly once by exactly one vehicle.
- All vehicle routes start and end at the depot.
- Every route has a total demand not exceeding the vehicle capacity.

The travel times between any two potential locations are given as input to the problem. Consequently, the total travel is computed by summing up the travel time over the chosen routes.

In reality, carrier companies are faced with a number of other issues not conveyed in the VRP. The research in this thesis introduces a number of realistic variants of the VRP. These variants consider the VRP as a core component and incorporate additional features. By definition the VRP is NP-hard. Throughout the years a vast amount of research was aimed at developing both exact and heuristic solution procedures. Building on this established literature, solution procedures are developed to fit the variants proposed in this thesis.

The standard VRP considers that the travel time between any pair of locations is constant throughout the day. However, congestion is present in most road networks. Considering traffic congestion results in time-dependent travel times, where the travel time between two location depends not only on the distance between them but also on the time of day one chooses to traverse this distance. Time-dependent travel times are considered in Chapters 2 and 3 of this thesis. Thus, in these Chapters we incorporate the *time* dimension into the VRP.

The standard VRP does not take into account any customer service aspect. The customers are presumed to be available to receive their goods upon arrival of the vehicles. However, a number of carrier companies quote their expected arrival time to their customers. We introduce the concept of self-imposed time windows (SITW). SITW reflect the fact that the carrier company decides on when to visit the customer and communicates this to the customer. Once a time window is quoted to a customer the carrier company strives to provide service within this time window. SITW differ from time windows in the widely studied VRP with time windows (VRPTW), as the latter are exogenous constraints. In Chapters 4 and 5 SITW are endogenous decisions in stochastic environments. Thus, in addition to the sequencings decisions required by the VRP further *timing* decisions are needed.

This thesis extends the VRP in two major dimensions: time-dependent travel times and self-imposed time windows. In reality carrier companies are faced with various uncertainties. The presented models incorporated some of these uncertainties by addressing three stochastic aspects: (I) In Chapter 3 stochastic service times are considered. (II) In Chapter 4, stochasticity in travel time is modeled to describes variability caused by random events such as car accidents or vehicle break down. (III) Finally, in Chapter 5 the objective was to construct a long term plan for providing consistent service to reoccurring customers. Stochasticity in this thesis is treated in an *a priori* manner. The plan, consisting of routes and timing decisions where necessary, is determined beforehand and is not modified according to the realization of the random events.

Chapter 2 addresses environmental concerns by studying CO_2 emissions in a time-dependent vehicle routing problem environment. In addition to the decisions required for the assignment and scheduling of customers to vehicles, the vehicle speed limit is considered. The emissions per kilometer as a function of speed, is a function with a unique minimum speed v^*. However, we show that limiting vehicle speed to this v^* might be sub-optimal, in terms of total emissions. We adapted a Tabu search procedure for the proposed model. Furthermore, upper and lower bounds on the total amount of emissions that may be saved are presented. Quantifying the trade-off between minimizing travel time as opposed to CO_2 emissions is an important contribution. Another important contribution lies in incorporating fuel costs in the optimization. As fuel costs are correlated with CO_2 emissions, Chapter 2 shows that even in today's cost structure limiting vehicle speeds is beneficial.

Chapter 3 defines the perturbed time-dependent VRP (P-TDVRP) model which is designed to handle unexpected delays at the various customer locations. A solution method that combines disruptions in a Tabu Search procedure is proposed. In Chapter 3 we identify situations capable of absorbing delays. i.e. where inserting a delay will lead to an increase in travel time that is less than the delay length itself. Based on this, assumptions with respect to the solution structure of P-TDVRP are formulated and validated. Furthermore, most experiments showed that the additional travel time required by the P-TDVRP, when compared to the travel time required by the TDVRP, was justified.

In Chapter 4 the notion of self imposed time windows is defined and embedded in the VRP-SITW model. The objective of this problem is to minimize delay costs (caused by late arrivals at customers) as well as traveling time. The problem is optimized under various disruptions in travel times. The basic mechanism of dealing with these disruptions is allocating time buffers throughout the routes. Thus, additional timing decisions are taken. The time buffers attempt to reduce potential damage of disruptions. The solution approach combines a linear programming model with a local search heuristic. In Chapter 4, two main types of experiments were conducted: one compares the VRP with VRP-SITW while the other compares VRPTW with VRP-SITW. The first set of experiments assessed the increase in operational costs caused by incorporating SITW in the VRP. The second set of experiments enabled evaluating the savings in operational costs by using flexible time windows, when compared to the VRPTW.

Chapter 5 extends the customer service dimension by considering the consistent vehicle routing problem. Consistency is defined by having the same driver visiting the same customers at roughly the same time. As such, two main dimensions of consistency are identified in the literature, driver- and temporal consistency. In Chapter 5, driver consistency is imposed by having the same driver visit the same customers. Furthermore, we impose temporal consistency by SITW. A stochastic programming formulation is presented for the consistent VRP with stochastic customers. An exact solution method is proposed by adapting the 0-1 integer L-shaped algorithm to the problem. The method was able to solve the majority of test instances to optimality.

About the author

Ola Jabali was born in Nazareth, Israel on April 21, 1980. She received her B.Sc. (2003) and M.Sc. (2006) in Industrial Engineering from the Technion - Israel Institute of Technology. Her master's thesis on resource Scheduling in Emergency Departments was supervised by the late David Sinreich. In parallel to her studies, she worked in industry and in a non-governmental organization.

In 2006, she started her PhD project on vehicle routing problems under the supervision of Tom van Woensel and Ton de Kok. During her PhD project she cooperated with Christophe Lecluyse and Herbert Peremans from the University of Antwerp, Belgium. Furthermore, she cooperated with Roel Leus from the Katholieke Universiteit Leuven, Belgium. Part of her PhD research was conducted at CIRRELT (Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation), Canada, in collaboration with Michel Gendreau and Walter Rei. On November 24 Ola defends her PhD thesis at Eindhoven University of Technology.