TIL5060 TIL THESIS

Time slot offering: the effect of various green labeling approaches on routing performance, considering customers' preferences

A case study at Crisp
Report number: 2023.TIL.8793

Author

M. CHAHBARI | 5365007

Supervisors

PROF. DR. R. R. NEGENBORN
Chair (3mE)

IR. M. B. DUINKERKEN
Daily supervisor (3mE)

DR. S. FAZI
Daily supervisor (TPM)

L. VERPLOEGH
External supervisor (Crisp)

April 28, 2023

Preface

This master thesis is written as the final research of the MSc Transport, Infrastructure and Logistics at the Technical University of Delft. The project was written at Crisp, an e-grocer in The Netherlands, which I started in September 2022. In this research, various green labeling approaches are analysed in order to improve relevant routing performance. Here, customers' preferences are included via choice modeling. Working towards more sustainable operations becomes increasingly important nowadays. Also for customers, who become more aware of their ecological footprint and therefore prefer to choose sustainable options more often. This subject was a good representation of a less-studied topic in the academic literature with high relevance in practice. I enjoyed working on my thesis a lot. I learned much from combining both wishes and requirements from TU Delft and Crisp and it was very informative to work with real data. Besides, this thesis really felt like a closure of my time as a student, where various research and design methods I learned over the past few years came together.

I would like to thank Mark Duinkerken, Stefano Fazi and Laurens Verploegh for their supervision, valuable feedback and flexibility and Rudy Negenborn for his critical and helpful suggestions. I would also like to thank Sander van Cranenburgh for his help with choice modeling. Furthermore, I would like to thank everybody at Crisp who helped me by giving me extra information. Lastly, I would like to thank my family and friends who supported me not only during my thesis but also for the rest of my career as a student. I hope you will enjoy reading this research and that it will contribute to a more sustainable last-mile delivery.

Summary

In recent years, business-to-consumer (B2C) e-commerce is growing rapidly. Due to the COVID-19 pandemic, the demand for online orders grew even more. New opportunities for different sectors emerge. One of them is the online groceries delivery service, whose revenue doubled between 2019 and 2021 in the Netherlands. The last-mile process can be seen as the most critical process in the chain which is responsible for approximately 50% of the costs in the whole chain. The reason for this is that the delivery to the end destination is the least efficient and most expensive part of the process due to demanding consumers in the aspect of delivery speed, time frames and other preferences. Unfortunately, most consumers are not willing to pay for these wishes, which has an impact on the revenue of the last-mile process. Furthermore, customers have in general the same preference for delivery windows, tightening the constraints of the routing model since demand is not evenly spread throughout the day.

Routing performance is mostly defined with indicators related to costs, while sustainability becomes an increasingly important topic. In order to improve the routing performance, most studies focus on improving the routing model. Another approach is to not focus on the model itself but on its most constraining inputs. For a routing model, the most constraining factors with regard to the input are how the time windows are offered, also called demand management. A relatively new concept within demand management for time window assignment is green labeling; time windows which contribute to improving the routing performance in terms of sustainability. Certain time slots are given a so-called green label. The main difference with the pricing or availability incentive is that environmental incentives are intrinsic motivators, while the former is an extrinsic motivator. It is challenging to design and compose the optimal set of time windows and incentives that contribute to improving the routing performance. The definition of a green label can be different. In this research, a time slot with a green label leads to better clustering, without limiting the number of options. This has as a result lower fuel consumption and emissions. From literature, limited information is known about how to implement green labeling and what the impact is on routing performance.

The main research question is:

How to implement green labeling for an e-grocer's time slot offering in order to improve relevant routing performances, considering customers' preferences?

With the following sub-questions:

1. How can a last-mile delivery system for an e-grocer be defined?
2. What are the different options with regard to time window offering and demand management?
3. How can a choice model be estimated in order to investigate the impact of green labeling on slot choice?
4. How can a simulation study be set up in order to study the effect of green label time window offering on the routing performance?
5. What is the effect of green label time window offering on the routing performance?

The first research question is answered by use of literature and desk research and has the goal to have a better understanding of current last-mile processes and components and how they are related.

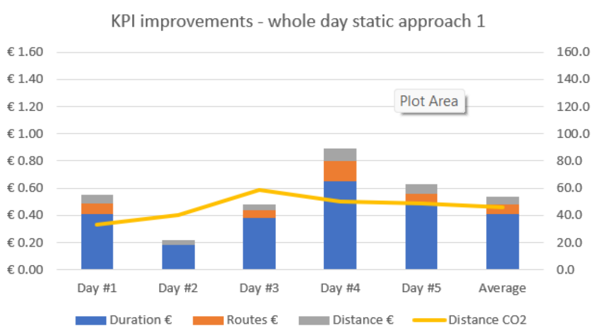
A last-mile delivery system consists of the following main components: pre-planning, order- and planning system, warehouses, hubs, delivery vehicles and the customer. Last-mile delivery entails the transport from the hub or warehouse to the customer. The process starts with forecast and pre-planning, which provides capacity constraints with regard to e.g. deliverers and vehicles. The order process is important in terms of demand management since the options that are given to the customer have a great effect on the rest of the process. Also, customer satisfaction is influenced. When the customer places an order, the planning process can start, followed by the picking process and the order can be delivered to the customer from the warehouse or via a hub. There are two KPIs when assessing routing performance, namely costs and sustainability, based on the factors duration, vehicles and distance, from which duration has the highest impact on costs and distance is the only factor impacting sustainability in this case study.

The second research question is also answered by use of literature and desk research and has the goal to analyse what is already known about demand management. Besides, time window offering in practice is analyzed in order to take learnings and inspiration from it. The standard classification of demand management concepts is subdivided by pricing or time slot allocation incentives, which can occur in a static or dynamic environment. In a static environment, the possible time slot options are determined before the booking horizon starts based on characteristics such as the delivery location and time of day. A dynamic approach has the goal to serve time offering options individually for every customer request that occurs during the booking horizon. A relatively new form of incentive is green labeling, meaning that the labeled time slot will contribute to lower fuel consumption and emissions. Besides a static or dynamic approach and used incentives, e-grocers can also differentiate their time window offering in terms of the number of time slots. Compared to other e-grocers, Crisp offers a wide variety of time slots and is, besides Picnic, the cheapest in terms of delivery costs. The time window offering in terms of window length and price occurs dynamically at AH and Jumbo. Crisp and Picnic deals with a static strategy, meaning that for a certain delivery location, the offering per day of the week is the same. Albert Heijn is the only e-grocer in the Netherlands that makes use of green-labeled time slots.

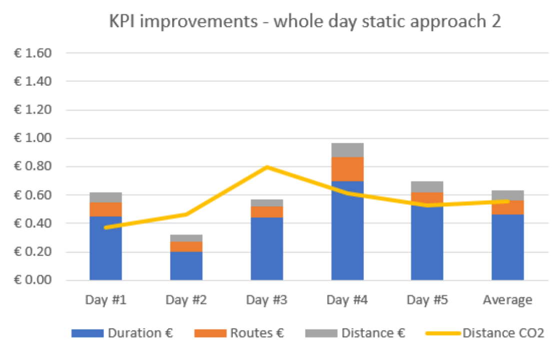
The third research question uses choice modeling and has the goal to determine how to set up a choice model and which attributes should be estimated. An MNL model will predict the probabilities of different options. First, different beta estimates are estimated in Apollo, a package in R, with the data of Crisp. These beta estimates are used in the utility function without the green label attribute in order to estimate the choice probabilities of the MNL model. This is done for the morning deliveries, the afternoon and evening deliveries together and for all the orders on the same day. Thereafter, the beta estimate of applying a green label is included. This has resulted in larger time slots being chosen more often. In total, there is an increase of 9.4% in large time windows.

The fourth research question uses choice simulation in order to design different static and dynamic time window offering approaches. In this research, the simulation study will be set up in a static and in a dynamic environment. In static time slot management (STSM), the larger time windows receive a green label. In here, two approaches are defined. In the first approach, all the four-hour time windows receive a green label. In the second approach, the four-hour time windows are changed to five-hour time windows. Here, there is accounted for that in for example low density areas, only evening deliveries are included. The green labeling will be done based on the zip codes of the accepted orders; the green labeling is based on their location if it is within a certain proximity of already accepted orders for that day. There will be two dynamic labeling approaches simulated within dynamic time slot management (DTSM). The first approach has almost the same rules for high-density areas compared to the static situation. A difference is that this approach only places a green label on the morning and afternoon four-hour time window in order to nudge customers more towards less popular day parts. For low-density areas, different rules will apply on which time window receives a green label. The second approach is similar, except that customers are not nudged more towards the morning or afternoon.

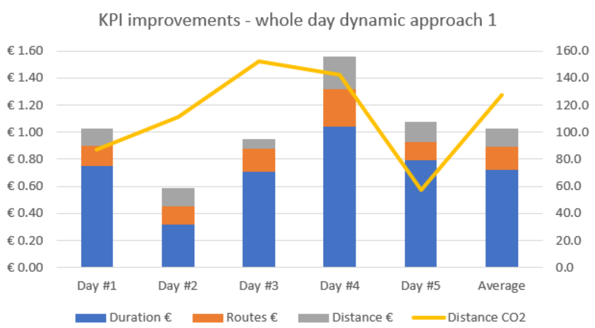
The fifth research question is answered by means of a route simulation study in order to investigate the effect of green labeling. Figure 1 shows the KPI improvements per order per approach, where the monetary factors are incorporated in the bars and the CO_2 factor is depicted by the yellow line.



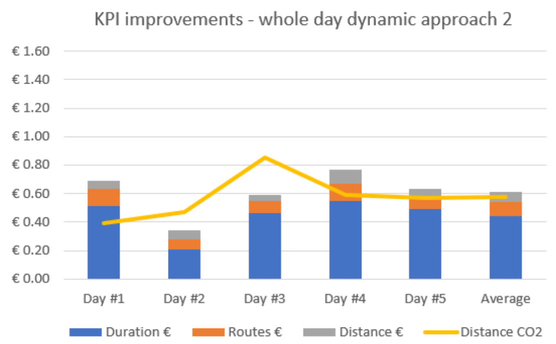
(a) Whole day static approach 1



(b) Whole day static approach 2



(c) Whole day dynamic approach 1



(d) Whole day dynamic approach 2

Figure 1: KPIs four different approaches

Table 1 shows a summary of the KPIs for the different approaches simulated for the whole day. *Dynamic - approach 1* is the approach that leads to the highest improvement in terms of costs and sustainability per order. This is followed by *Dynamic - approach 2* that have better improvements in terms of sustainability compared to the number three, *Static - approach 2*. In terms of costs, number three performs slightly better than number two, but an improvement in sustainability weighs more. The lowest in the ranking is *Static - approach 1*.

	Costs	**Sustainability**	**Ranking**
Static - approach 1	€0.53	46.2 $g\ CO_2/km$	4
Static - approach 2	€0.63	55.4 $g\ CO_2/km$	3
Dynamic - approach 1	€1.03	127.4 $g\ CO_2/km$	1
Dynamic - approach 2	€0.61	57.5 $g\ CO_2/km$	2

Table 1: Summary of the KPIs improvements per order of all the approaches

It can be concluded that dynamic green labeling has the most promising effect on routing performance, especially when customers are more nudged towards the largest time windows in the less popular day parts.

The choice modeling and simulation and route optimization are both based on some assumptions. The results are not a full reflection of reality. The choice model is adapted to the time window offering of Crisp, which has the benefit that it reflects better a real process. A downside is that it is less generic. Also, only attributes with regard to the time window itself are included and not characteristics linked to the customer. With regard to route optimization, only one route software package is used and comparison between outcomes is therefore not possible. Also, only four approaches are tested, while there are countless options possible.

The research gap is reduced since this research contributed to increasing the knowledge on how to implement green labeling and what its effect is on relevant routing performance. Data from an e-grocer is used and a choice model was included. This combination was not used in scientific research before. Recommendations on a practical level include that testing in real-life is needed in order to determine the real preferences with regard to green-labeling. Also, discussion with route software companies should clarify which steps should be taken in order to be able to implement dynamic green-labeling in the software. It is recommended to first implement a static approach before implementing a dynamic one. In this way, learning from the static one can be taken into account. On a more scientific level, it is recommended to do more research on other options with regard to distance determination for the grouping. Now only zip codes are used. It is also recommended to do more research on choice modeling within this context and to analyse what the effect is of different attribute values and customer-specific attributes.

Contents

List of abbreviations

- **AHD** = attended home delivery
- **MILP** = mixed integer linear programming
- **MNL** = multinomial logit
- **GAM** = general attraction model
- **KPI** = key performance indicator
- **BPMN** = business process model notation
- **STSM** = static time slot management
- **DTSM** = dynamic time slot managment
- **VRP** = vehicle routing problem

List of Figures

List of Tables

1 Introduction

1.1 Problem definition

In recent years, business-to-consumer (B2C) e-commerce is growing rapidly (Mangiaracina et al., 2019). Due to the COVID-19 pandemic, the demand for online orders grew even more. New opportunities for different sectors emerge. One of them is the online groceries delivery service, whose revenue doubled between 2019 and 2021 in the Netherlands (Chevalier, 2021). The online grocery share with regard to revenue is estimated at 5.6% of the total supermarket revenue in the Netherlands in 2022, which is responsible for about 2.5 billion euros in revenue. This includes only households. In 2019, the market share of only groceries was 3.2%, indicating the rapid growth of this sector (van Loon, 2022). The last-mile process can be seen as the most critical process in the chain which is responsible for approximately 50% of the costs in the whole chain (Mangiaracina et al., 2019). The reason for this is that the delivery to the end destination is the least efficient and most expensive part of the process due to demanding consumers with regard to delivery speed, time frames and other preferences. Unfortunately, most consumers are not willing to pay for these wishes, which has an impact on the revenue of the last-mile process. Furthermore, customers can have in general the same preference for certain delivery windows, tightening the constraints of the routing model since demand is not evenly spread throughout the day. Next to that, there are challenging targets with regard to service levels, small dimensions of orders and the high level of distribution of destinations (Macioszek, 2017).

It is important to improve the routing performance in order to increase the efficiency of the whole delivery chain. Especially in the e-grocer sector, where shorter time windows than the ones in, for example, the parcel delivery sector, are used because of the necessary attended home delivery due to the nature of the goods. Routing performance is mostly defined with indicators related to costs, while sustainability becomes an increasingly important topic. In order to improve the routing performance, most studies focus on improving the routing model. Another approach is to not focus on the model itself but on its most constraining inputs. For a routing model, the most constraining factors with regard to the input are how the time windows are offered, also called demand management (Köhler et al., 2020). A relatively new concept within demand management for time window assignment is green labeling; time windows which contribute to improving the routing performance in terms of sustainability. Certain time slots are given a so-called green label. The main difference with the pricing or availability incentive is that environmental incentives are intrinsic motivators, while the former is an extrinsic motivator. It is challenging to design and compose the optimal set of time windows and incentives that contribute to improving the routing performance. The definition of a green label can be different. An option can be that a time slot with a green label leads to better clustering, without limiting the number of options. This has as a result lower fuel consumption and emissions (Agatz et al., 2021). Another option can be that for that time slot, an electric vehicle will be used. From literature, limited information is known about the choice preference with regard to green labeling and its impact on routing performance.

1.2 Research objective

In every sector, and also in the e-grocer sector, sustainability becomes a more important topic. Especially in the last-mile, which is the most inefficient process in terms of costs and sustainability. There is constantly the need for research on how to improve time window offering in order to optimize routing performance. There must always be a balance between the flexibility and amount of time slots offered and customer satisfaction; less choice leads to better routing performance but worse customer satisfaction. In this context, it is needed to do more research on the effect of green labeling, a new concept within the e-grocer sector, on routing performance and which scenario fits best. Here, customers' preferences should be taken into account in order to guarantee the balance between the optimal time window offering and customer satisfaction. Also, in transport, improving routing performance covers reducing total distance, hence costs and CO_2 emissions. This research uses data from Crisp, an online e-grocer in the Netherlands, but this is also applicable to other e-grocers or even comparable sectors outside grocery.

1.3 Research questions

The sub-questions will be set up in such a way that answering them all will result in answering the main question. This leads to achieving the research objective as described in subsection 1.2. The main research question is:

How to implement green labeling for an e-grocer's time slot offering in order to improve relevant routing performances, considering customers' preferences?

With the following sub-questions:

1. How can a last-mile delivery system for an e-grocer be defined?
2. What are the different options with regard to time window offering and demand management?
3. How can a choice model be estimated in order to investigate the impact of green labeling on slot choice?
4. How can a simulation study be set up in order to study the effect of green label time window offering on the routing performance?
5. What is the effect of green label time window offering on the routing performance?

1.4 Methods

The different research methods are described in detail in section 3. For the first two research questions, both literature research and desk research are used. For the first research question, also interviews with experts are conducted. With this information, a choice model is set up. The choice model is estimated with real data and evaluates the preferences of customers. The output from the choice model will be used as input for the optimization of route performance. This will be done in Routigo, a route-planning software package.

1.5 Scope of this research

An important note is that this research focuses on the effect of green labeling on the routing performance of an e-grocer, including customers' preferences. Some assumptions on the beta estimates of green labeling are made since no survey on choice behavior is conducted in this research. Since the routing software of Crisp is used, the route planning takes place in a so-called 'black-box'. Settings of the software are described so that replication of the simulations is possible. An own developed vehicle routing model is therefore out of scope. In order to determine which customers receive which green labels on which time windows, different approaches are described. The clustering will take place on zip code level, since integrating a distance matrix is out of scope of this research because external routing software is used.

1.6 Structure of report

In order to achieve the objective of this research, the approach visualized in Figure 2 will be followed. After this section 1, the introduction, a literature review is conducted in order to have an understanding of what is already known and studied within relevant fields, which can be found in section 2. This is followed by the methodology in section 3 where the research methods and framework are described in detail. In section 4, the system description, a last-mile delivery system for an e-grocer will be described, together with the different time window offering methods and demand management. The first and second research questions will be described here. The third research question will be answered in section 5. Here, a choice model is designed and executed in order to estimate the choice behavior with regard to time window offering. In section 6, a simulation study will be set up in order to study the effect of green label time window offering on routing performance, the fourth research question. The effect of green labeling will be discussed in the last sub-question in section 7. Here, the different scenarios will be compared to the base case of an e-grocer. In section 8, the results are discussed, followed by the conclusion and recommendation in section 9, which answers the main research question.

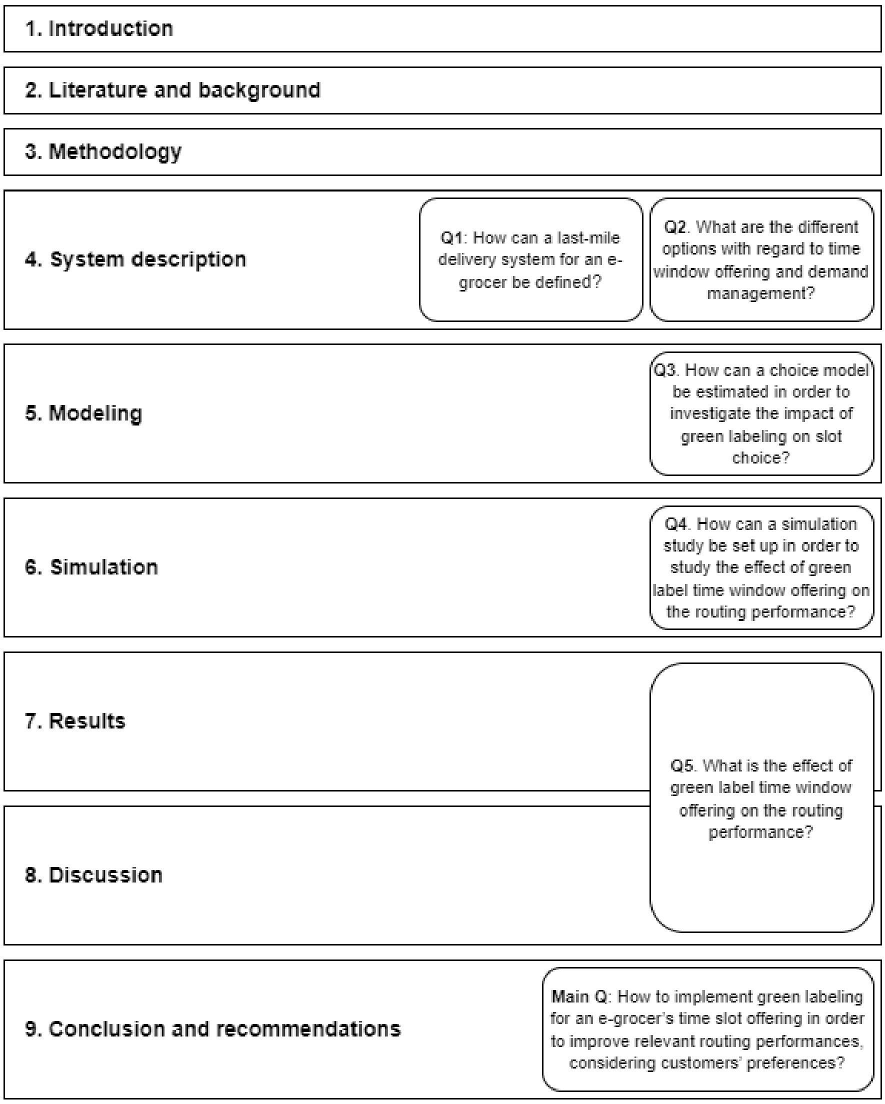


Figure 2: Structure of this research

2 Literature and background

Last-mile can be defined as the final stretch of the whole supply chain. The supply chain starts with a manufacturer. This can be considered as the first part of the whole chain, also called the first-mile. The products are stored in a distribution center and from there, the products are transported by vans or trucks to a hub or warehouse. Here, the middle-mile starts and the products are transported to other hubs or warehouses, which can be abroad. From there, vans or trucks will transport the products to the warehouses where the last-mile will start. Lastly, vans will deliver the products to the customer (Truong, 2022). This whole process is shown in Figure 3. Note that deviations are possible, depending on the size and other characteristics of the chain.

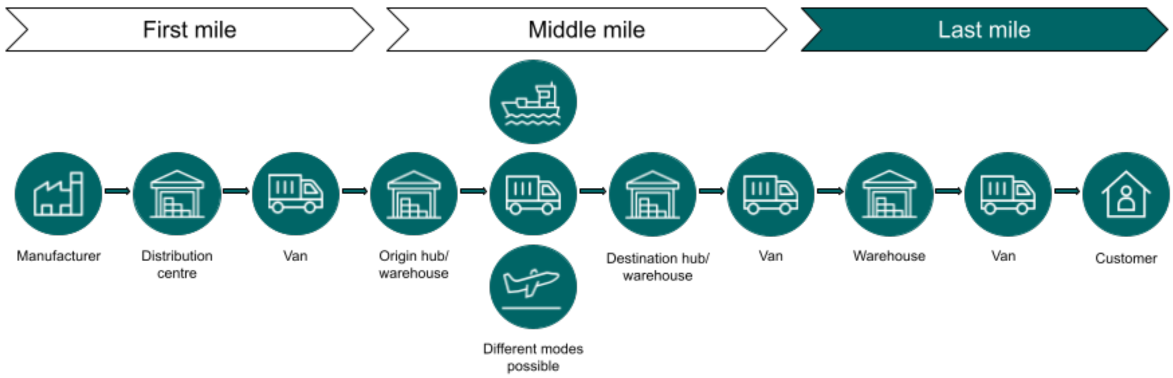


Figure 3: Standard supply chain process

This chapter includes relevant background which is needed to better understand the context of the problem itself and the e-grocer sector in the Netherlands. This is followed by a literature study on choice modeling and time slot offering.

2.1 The e-grocer model

The online grocery market is growing rapidly. In 2017, the total revenue of the market was around 500 million euros in the Netherlands. Five years later, the total revenue is estimated around 2.5 billion euros (Chevalier, 2021). The player with the biggest market share is Albert Heijn, followed by Picnic and Jumbo, who together have around 95% of the market (Insider, 2021). The online-only e-grocer model has several advantages compared to the traditional grocery model. Their inventory does not have to be distributed around a lot of locations, but only a few warehouses or hubs. This has as a result less spoilage and an increased inventory turnover rate which makes it possible to sell even better perishable products. Another advantage is that very detailed information about the buying behavior of the customers can be gathered, without using a loyalty card (Hays et al., 2005). Other advantages of online supermarkets can be lower prices due to no or less costs of brick-and-mortar stores, convenience and time savings (Hays et al., 2005). Next to that, the focus lies on people with a busy lifestyle or for whom doing groceries is difficult, for example, people with physical disabilities or the elderly. Targeting specific groups can create a competitive advantage. There are also some disadvantages of the e-grocer model. Brick-and-mortar stores have prominent locations, brand names and a large customer base. Customers like to pick their fruit and vegetables themselves and appreciate the contacts with the local bakery. The largest challenge for online supermarkets is order fulfillment and home delivery. The higher the deliveries in a certain area, the lower the costs per delivery and the higher the deliveries per hour. This has also a high impact on the sustainability performance of the deliveries (Hays et al., 2005).

The delivery of orders can be done in multiple ways. Common ways of delivery are attended home delivery (AHD), unattended delivery and in-store pickup (Mackert, 2019). This research focuses on AHD. When placing an online order, customers can choose a time window. For e-grocers, it is an important task to assign and meet the delivery time windows. A failed delivery is costly since the products are perishable. The offering and assignment of time slots is called demand management. The e-grocer can assign the time windows in a static or dynamic way. In a static approach, optimization only occurs after the cut-off moment while the dynamic approach also optimizes before the cut-off moment, for example after every order. Next to that, the time slot offering can be characterized by which time slot is offered throughout the day and the prices of the time slots.

2.1.1 The e-grocer model in practice

In the Netherlands, Picnic is the biggest e-grocer in terms of business-to-consumer orders, with a market share of 38%, followed by Albert Heijn with 37%. Jumbo has a share of 20%, leaving 5% share left for the other e-grocers, including Crisp. The growth of the three biggest e-grocers over the last seven years is shown in Figure 4.

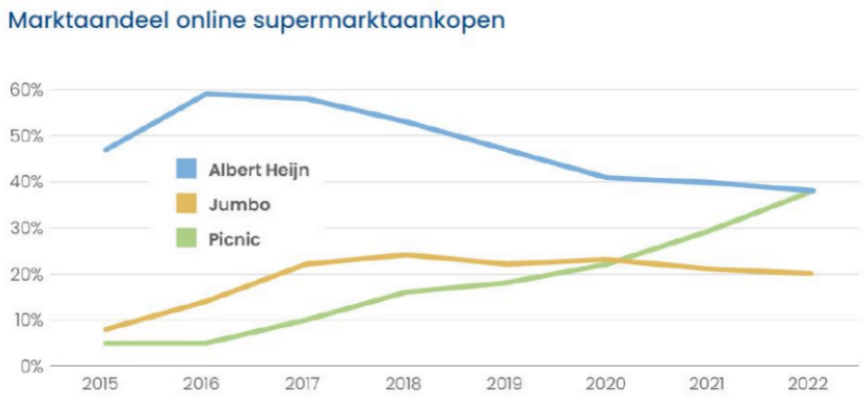


Figure 4: Market share among the biggest e-grocers (van Loon, 2022)

With regard to market positioning, the above-mentioned e-grocers are different. According to Treacy and Wiersema (1993), there are three main disciplines; product leadership, customer intimacy and operational excellence. It is important to master them all but to excel in one. Product leadership means excelling in product quality, innovation and brand marketing. Concerning customer intimacy, it is important to have a high value on attention for customers. Products and services are adjusted according to the wishes of the customers. Customer relations management (CRM) is therefore very important. Delivery of products and services must be on time. For operational excellence, it is of high importance to have an efficient production process. The focus is on high margins and products or services which are streamlined and automated.

When focusing on the e-grocer sector in the Netherlands, Albert Heijn and Jumbo can be classified as companies that excel in product leadership and Picnic excels in operational excellence. Crisp has a high value on customer intimacy but also on product leadership. Feedback from the customers is of great importance. Furthermore, the app and the delivery process should be one of the best. An overview is shown in Figure 5.

Figure 5: Marketing disciplines according to Treacy and Wiersema (1993)

Companies create time windows that are mostly differentiated by price or availability. This differentiation links to nudging customers towards more preferred time windows, which can be longer time windows, time windows that are not during peak hours, or both, which contributed to improving the routing performance. This is a difficult topic since on one side a company wants to increase its customer satisfaction by offering a lot of time windows against low prices or even for free, while on the other hand, an optimal routing performance is desirable. One major drawback is that price incentives have a high impact on the margins. Especially within the e-grocer sector, where the margins are already low. Another downside is that dynamic pricing can be seen as unfair from the point of view of the customer, which has a negative impact on customer satisfaction. With regard to availability, the time slot options can be different based on the location of the customer in order to cluster the customers and reduce the driven kilometers. Since this leads to fewer options for the customer, customer satisfaction is again negatively impacted (Agatz et al., 2021).

2.2 Literature review on demand management and choice modeling

A lot of research is done on combining demand management, also called nudging, with time slot allocation for attended home delivery (AHD). The most related literature on slotting policies in an AHD environment is classified in terms of decision and modeling characteristics in Table 2. The first column shows which *choice model* is used: fixed probability, multinomial logit (MNL) or general attraction model (GAM). *Demand management* can be defined as the optimization of actions that have an effect on demand, for example, pricing, availability or sustainability incentive. With *optimization algorithm* is meant which algorithm or model is used in order to optimize the route planning. *Cost estimation* describes how is determined what the costs of a new order are and therefore input to the optimization algorithm. The last column states the *slotting policy*, which can be static or dynamic. This research uses an MNL choice model and a mixed integer linear programming (MILP) model for optimizing from an external routing software package, which is described in section 5. In the next subsections, the literature will be reviewed on the choice models fixed probability, multinomial logit (MNL) and general attraction model (GAM).

Study	Choice model	Demand management	Optimization algorithm	Cost estimation	Slotting policy
(Campbell and Savelsbergh, 2005)	Fixed probability	Availability	-	Insertion heuristics	Static & dynamic
(Campbell and Savelsbergh, 2006)	Fixed probability	Pricing	-	Insertion heuristics	Static
(Agatz et al., 2011)	-	Availability	-	Continuous approximation	Dynamic
(Yang et al., 2016)	MNL	Pricing	-	Insertion heuristics	Dynamic
(Yang and Strauss, 2017)	MNL	Pricing	Approximating dynamic program	Continuous Approximation	Dynamic
(Klein et al., 2019)	MNL	Pricing	Approximating dynamic program	MILP	Dynamic
(Mackert, 2019)	GAM	Availability	MILP	Continuous approximation	Dynamic
(Agatz et al., 2021)	MNL	Green label	Simulation	Insertion heuristics	Static & dynamic
(Strauss et al., 2021)	MNL	Pricing	Linear program	Continuous approximation	Dynamic
This research	MNL	Green label	MILP	-	Static & dynamic

Table 2: Classification of related literature on demand management and choice modeling in AHD

2.2.1 Fixed probability

According to a fixed probability model, the chance that a person would select a specific alternative from a list of possibilities is determined by their preferences, which are considered to be fixed for each person. The model explains how the probability of selecting each alternative relates to the qualities of the alternatives (such as price, features, or quality). The choice model based on fixed probabilities, for instance, implies that a customer has a fixed set of preferences that define the likelihood of selecting each product when the consumer is offered multiple alternative items. These preferences may be influenced by the product's characteristics, brand reputation, price, or quality. The model can be used to calculate the likelihood that a customer will select each option (Campbell and Savelsbergh, 2005).

In the paper of Campbell and Savelsbergh (2005), all customers have a certain slot profile containing the slots that the customers are willing to accept. They are among the first to study customer requests via dynamic slotting. The demand is represented as an arrival process that is not affected by the company's decisions. After every customer request, the offered time windows are adjusted. A simplistic choice model is used based on fixed probability. In their subsequent work, Campbell and Savelsbergh (2006) model the time slot decisions of the customers influenced by price incentives. The incentive decisions are based on the use of insertion costs of a customer request based on already accepted customers. Only the marginal delivery costs are taken into account and the crucial displacement costs for future orders are neglected. Also here, a simplistic choice model is used, this time influenced by price.

The advantages are that the simple estimation and interpretation of fixed probability models make them accessible to researchers and analysts without specialized statistical knowledge. A downside of fixed probability is that the assumption is made that people have fixed preferences that do not alter with time or in response to new information, which may not always be the case in reality. Fixed probability models are more concerned with forecasting choices than with clarifying the reasons behind people's decisions, resulting in lower explanatory power. Another advantage is that the model has a high computational efficiency. These models are also dependent on data quality. The correctness of fixed probability models depends on the accuracy of the data used to estimate them, especially the accuracy of the metrics used to represent the attributes of the alternatives being presented.

2.2.2 Multinomial logit model

A multinomial logit (MNL) model is a statistical tool for the analysis of discrete choice data, which refers to situations where a person must select one option from a range of potential options. In an MNL model, the likelihood that a person would select a certain alternative can be modeled as a function of both the qualities of the person and the properties of the alternatives. The MNL model presupposes that the decision-maker chooses the alternative that, given its properties and personal characteristics, has the highest utility. With coefficients denoting each attribute's relative weight, the utility function is assumed to be linear in the alternative attributes. The MNL model also presupposes that the error term is equally and independently distributed among people and possible solutions (Yang and Strauss, 2017).

Yang et al. (2016) considers a dynamic pricing problem. An MNL choice model is estimated with real data, which is used for the continuous calculation of opportunity costs. The demand is very sensitive to delivery prices and time slot availability. With regard to pricing, only the opportunity cost estimates based on marginal routing costs are used. The effect of potential future lost revenues is not included. The research of Yang and Strauss (2017) is similar, only by using approximate dynamic programming the potential losses are included. An alternative way to approximate opportunity costs by approximate dynamic programming is used which potentially better incorporates displacement costs. This study is computationally more efficient than Yang et al. (2016), but the performance with regard to profit is worse. Klein et al. (2019) uses a mixed integer linear program (MILP) with the dynamic pricing model of Yang et al. (2016) for the cost estimation. Comparable to Agatz et al. (2011), a seed-based scheme is used. A downside is that the MILP is computationally inefficient due to a large number of decision variables. The work of Agatz et al. (2021) is the first to study choice behavior within green-labeled incentives for attended home delivery. Insertion heuristics are used to determine the costs of every new customer arrival. The potential impact of steering demand is quantified by the use of simulation. Strauss et al. (2021) combines the concepts of demand management and vehicle routing, as well as flexible products. Customers can choose one or multiple time slots. A tractable linear programming formulation links demand management decisions and route cost implications while including choice behaviour. A dynamic pricing policy on flexible time slots is the result. In all the above cases, an MNL choice model is used.

Some advantages of using an MNL model are its flexibility; more than two options can be considered and also choices that have different constraints. Besides, the resulting coefficients are easy to interpret since it shows the marginal utility of each attribute. A disadvantage is that the MNL model makes the assumption that each alternative's decision probabilities are independent of those of the other alternatives. In real-world circumstances, the availability of other options may impact a person's decision, which goes against this supposition.

2.2.3 General attraction model

A general attraction model (GAM) tries to explain the decisions that are made when faced with a variety of alternatives. According to a GAM, individuals are attracted to different attributes, for example price or time. The attributes have different weights for different individuals. A GAM is a type of random utility model (Mackert, 2019).

Mackert (2019) takes into account that a customer request is only accepted if its generated revenue is high enough. The probability that a customer chooses a different time slot if the preferred one is not available is included. A generalized attraction model (GAM) is used, which overestimates less than an MNL model since dissatisfaction is included. The cost approximation is based on the seed-scheme of Klein et al. (2019). Using a GAM in this manner will not be usable in solving the problem of this study, since information about potential revenue is needed.

Also, a GAM has its flexibility as an advantage. Complex interactions can be captured between attributes of the different alternatives. Another advantage is its ability to estimate heterogeneity. Since it can estimate the different weights that individuals have for different attributes, individual preferences are incorporated better. A downside is its complexity and therefore lower computational efficiency. A large amount of data and computational power is needed.

3 Methodology

3.1 Research methods

An overview of the different methods per sub-question is given in Table 3 with a description of the methods below. In total, five different research methods are used.

Questions	Goal	Methodology	Chapter
1. How can a last-mile delivery system for an e-grocer be defined?	Have a better understanding of current last-mile processes and components and how they are related.	Literature research Desk research Interviews with experts	section 4
2. What are the different options with regard to time window offering and demand management?	To analyse what is already known about demand management, how competitors have organized their demand management and to take learnings and/or inspiration from this.	Literature research Desk research	section 4
3. How can a choice model be estimated in order to investigate the impact of green labeling on slot choice?	Determine how to set up a choice model and which attributes should be included. Different attributes are estimated.	Choice modeling	section 5
4. How can a simulation study be set up in order to study the effect of green label time window offering on the routing performance?	Design different approaches where static and dynamic time window offering is included based on the choice model. Define the working and settings of the routing optimization model.	Choice simulation	section 6
5. What is the effect of green label time window offering on the routing performance?	Run route optimizations in order to investigate what the effect is of green labeling. Crisp's current routing software Routigo will be used. Resulting KPIs are compared to historic KPIs.	Optimization study	section 7

Table 3: Sub questions and their corresponding methods

Literature research and desk research
In order to answer questions 1 and 2, mainly literature research and desk research will be used. In order to understand current processes within last-mile delivery better, literature research can be of great importance. A lot can be found on demand management and choice models, as can be seen in Table 2. Since some information about competitors cannot be found in literature but more on websites, news articles and blogs, desk research will also be used.

Interview with experts
Experts in the field of last-mile delivery will be interviewed in order to receive some additional information that can be used for the literature and desk research or as an addition to the findings. This can be employees at an e-grocer or at other relevant companies.

Choice modelling
In order to investigate the impact of green labeling on slot choice, a choice model will be used. Which exact choice model will be used will be discussed in section 5. A requirement analysis will be performed in order to determine which choice model fits best.

Choice simulation
When from choice modeling becomes clear what the beta estimates are of the different attributes, multiple simulation approaches have to be described in order to perform the route simulation studies later on. The choice simulation describes which time window offer each customer receives and what the impact is on the customer that follows. Time window offering can be based on location, previous customers and order size.

Optimization study
Companies mostly make use of external optimization software in order to optimize route planning. The output from the discrete choice analysis will be used in order to adjust the order data of the case study. The historic route planning will be compared to the route planning from the simulations where the chosen time windows are adjusted due to green labeling.

3.2 Research framework

In this research, the main research methods are choice modeling and route optimization. A choice model will be designed which is able to calculate certain characteristics of attributes. In order to do so, order data is needed, together with the beta estimate related to green labeling. With this information, it is possible to estimate choice behavior and to determine other relevant beta estimates resulting from the order data. Information from the system description is needed in order to understand relevant processes. The choice model is the foundation of the third research question. The choice model will also be used in order to simulate choice behavior with regard to green labeling. Again, the original order data is needed. The next step is to adjust orders in the order database and the new order data can be inserted into the route planning software for route optimization. This will all be treated in the fourth research question. Lastly, the last research question will be treated in the results section. An overview is shown in Figure 6.

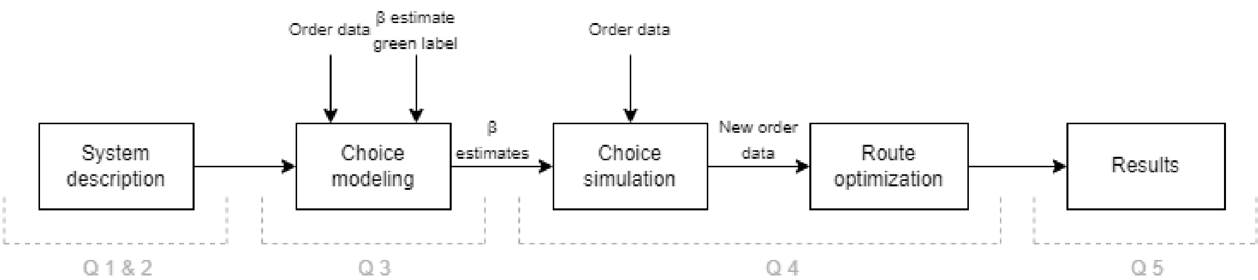


Figure 6: Research framework

4 System description

This chapter will perform a system description where the last-mile delivery system will be described and time window offering and demand management. This is the first part of the research framework as shown in Figure 6. The last-mile delivery system for an e-grocer will be analyzed by means of defining the different system components and their relation to each other. Also, the key performance indicators (KPIs) will be described for the last-mile system. For the second research question, different time window offering methods from literature and from practice are described. This is done in order to have a better understanding of performance and innovations within time window offering in the e-grocer sector. Differences in time window offering and demand management will be analyzed.

The following sub-questions will be answered in this section:

Sub question 1: How can a last-mile delivery system for an e-grocer be defined?

Sub question 2: What are the different options with regard to time window offering and demand management?

4.1 Last-mile delivery system description

Simply formulated, the last-mile delivery system entails the transport from the hub or warehouse to the customer. In parcel delivery, there can be a pick-up point or a locker before the parcel arrives at the end consumer but for AHD, this is not possible. The main physical process within last-mile delivery is the transport from hubs or warehouses to the end consumer. This is mostly done via delivery vans, which can be combustion based or electric. Mainly in city centers, other options are also possible. Examples are electric bikes and light electric vehicles. When the amount of orders and market area increases, hubs are needed in order to deliver to all the customers. When introducing hubs, transportation to the hubs from the warehouse becomes a new and important process. Trucks have to be implemented in order to supply the hubs. Here, there are two options with regard to route preparation which influences last-mile processes; it is possible to either transport the parcels not on route order and move the sequencing process to the hubs or to put for every route the parcels on a roll container and keep the sequencing process in the warehouse. A downside of the latter option is that the optimal capacity of the truck is not used. An advantage is that the sequencing process can be made more efficient due to volume. Other important, less visible, processes are the order and route planning systems. The design of these systems determines a large part of the performance of the operations. Customers are influenced by how time windows and other relevant information and options are presented to them. Also, delivery costs play an important role.

One reason why the last-mile process is the least efficient has to do with economies of scale occurring at different stages. First of all, it is more efficient for manufacturers to produce products in large quantities in order to lower costs. Therefore it can be more efficient for parties later in the chain to buy products from those large manufacturers. With regard to transport, at the beginning of the chain, larger quantities can be transported simultaneously. This makes it possible to fill large trucks or even ships or planes. This makes it more convenient to transport internationally. Lower in the chain, it becomes more difficult to gain from those economies of scale. Since the quantities to transport become smaller, the transport costs will increase. This is the main reason why the last-mile process is the most expensive and inefficient part of the whole supply chain.

An overview of the last-mile system is shown in the Business Process Model and Notation (BPMN) scheme in Figure 7. Here, the last-mile process starts at the e-grocer side with a sales forecast. A forecast is very important for later stages since it determines the capacity. This is followed by a pre-planning where is indicated how many deliverers and vehicles are needed. The customer starts the process by placing an order, then the e-grocer processes the order and the planning can be made. The customer receives an estimated time of delivery and the e-grocer can continue with the picking process. There are two possible options after picking; either the delivery goes straight to the client or the delivery will be first sent to a hub with a truck and then delivered to the client.

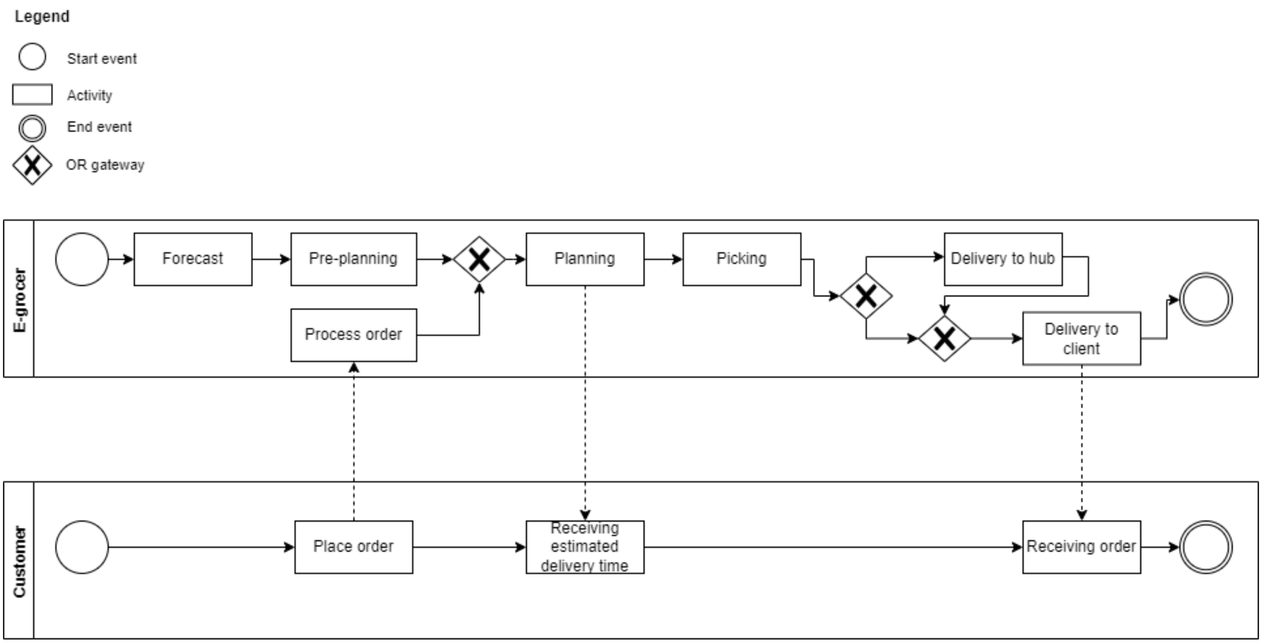


Figure 7: Process of a general last-mile system

4.1.1 Last-mile system components

This section will explain the definition of the different components of the last-mile system and how they are related to each other. Each component can be found in Figure 7. Also, which in- and output are needed for each component is described.

Customer

The choice that is made by the customer has a great impact on the performance of the company. The customer starts the order process and also ends it. As discussed earlier, the company can nudge the customer towards certain options but it will always impact customer satisfaction. It is very important for a company to know what the wishes and preferences are of the customer and to deal with this in a smart and efficient way. The customer is becoming more flexible, works more from home and has an increasing interest in sustainable delivery (Agatz et al., 2021). Since the customer is able to give negative reviews, it is even more important to value customer satisfaction.

Pre-planning

Before the planning can be performed, it is important to perform a pre-planning based on the forecasted sales. With this information, it can be estimated how many vans and deliverers are needed and what the impact on the picking and routing process will be. The forecasted sales determined how many deliverers are scheduled for the corresponding days and therefore determine how many routes are possible. The number of available vans serves as the maximum for the number of routes. The division between normal vans and electric vans is also of great importance.

Order system
The order system can take different forms. Customers will place their orders mostly via a website or an app. Orders can be placed throughout the day. For next-day delivery, a certain cut-off moment is scheduled. Customers can see all the products, mostly clustered per category. Next to that, order systems are nowadays made smarter and more interactive in order to give the customer a better shopping experience. Features as 'recommended for you' is a typical extra element in order systems. Also, when a product is out of stock, it can be possible that similar products will be shown to the customer. After the customer placed an order, a few updates will be given to the customer which can be an estimation of the delivery time and a track-and-trace. Another important feature of the order system is that it has a great influence on route planning because of the time window offering. Time window offering can be a tool to cluster customers in certain areas but this also has an impact on customer satisfaction. Smaller time windows have a constraining effect on route planning but increase customer satisfaction.

Planning system
Another important system is the planning system. After the cut-off moment for the next day, the planning system will plan all the orders based on certain settings and inform the customer. Most companies make use of planning software from an external company. It is possible that the planning software can also be used for the hubs to track the delivery vans in real-time and to respond to disruptions.

Warehouse/hub
A warehouse can be described as a physical building to store goods. In the e-grocer sector, it is sometimes called a fulfillment center since the picking of products is an important process here. Suppliers deliver their goods to the warehouse and from there, the picked orders will be transported to either a hub or directly to a customer. When a hub is involved, trucks need to transport goods and/or picked boxes to the hubs. This is an extra element in the last-mile process. Another important role of warehouses and/or hubs is that the delivery vans are stored here. Therefore, a lot of parking space is needed and what is also an important new development is the charging of the vans. Charging poles are nowadays necessary for warehouses and/or hubs.

Delivery vehicle
The final part of the last-mile delivery is the physical delivery of the products. A delivery vehicle receives its route mostly via an app integrated with a navigation system. Sometimes, the deliverer also takes in bottles with a deposit or cardboard. Certain aspects of the vehicle are very important for planning purposes, such as the volume capacity and range for electric vehicles. Besides, the deliverer has another important role since this is the only physical contact that the customer has with the company.

4.2 Key performance indicators

There are different ways to calculate the routing performance of a planning system. The key performance indicators (KPIs) are costs and sustainability. With regard to lowering costs, duration contributes the most, followed by distance. The number of routes needed to plan, and therefore the number of vehicles needed, contribute the least. When performing a simulation for a whole day, a vehicle can be used multiple times. This is left out of scope. Next to costs, another KPI is sustainability. When implementing green labels, it is the most important factor. Decreasing the total distance contributed most to increasing sustainability since it is directly linked to polluting emissions or energy usage. Costs can be expressed in euros, while sustainability can be expressed in emissions.

For *duration*, every *hour* that a deliverer is needed costs around €40 according to Crisp. This includes loans, potential damages and sickness. If fewer routes are needed, fewer *vehicles* are needed. It is calculated that a lease transporter costs €1000 per month, therefore €33.33 per day. For distance, the fuel costs per *kilometer* are estimated to be €0.17. This results in the costs formula shown in Table 4. The average CO_2 emissions per kilometer for a delivery van is estimated at 158.4 grams (Williams, 2022). For simplicity, it is assumed that all deliveries are performed by a diesel van.

KPI	**Formula**	**Unit**
Costs	$40 \cdot$ hours $+ 33.33 \cdot$ vehicles $+ 0.17 \cdot$ kilometers	€
Sustainability	$158.4 \cdot$ kilometers	$g\ CO_2$

Table 4: KPIs with regard to routing performance

Sub question 1: How can a last-mile delivery system for an e-grocer be defined?

A last-mile delivery system for an e-grocer differs from a general last-mile delivery system. Due to attended home delivery, it is important to prevent delivery failures due the perishable nature of the delivery goods. Delivery via a pick-up point is not possible. A last-mile delivery system is the last part of the whole supply chain and consists of the following main components: pre-planning-, order- and planning system, warehouses, hubs, delivery vehicles and the customer. Last-mile delivery entails the transport from the warehouse or hub to the customer. The process starts with forecast and pre-planning, which provides capacity constraints with regard to e.g. deliverers and vehicles. The order process is important in terms of demand management since the options that are given to the customer have a great effect on the rest of the process. Also, customer satisfaction is influenced. When the customer placed an order, the planning process can start, followed by the picking process and the order can be delivered to the customer from the warehouse or via a hub.

There are two KPIs when assessing routing performance, namely costs and sustainability, based on the factors duration, vehicles and distance, from which duration has the highest impact on costs and distance is the only factor impacting sustainability.

4.3 Time window offering and demand management

The e-grocer sector is characterized by the requirement of attended home delivery (AHD). Due to the need for AHD, working with time windows is of high importance. Working with a larger time window length decreases the routing costs due to more efficient planning. What the effect is on other routing performance indicators, such as sustainability, is studied less. All e-grocers face the trade-off between minimizing routing costs and maximizing customer satisfaction, of which the latter is positively influenced by narrow delivery time windows (Yang et al., 2016). Next to that, the e-grocer sector is characterized by thin profit margins and a highly and fast-changing competitive market. Therefore, it is of great importance to manage the time window offering (Mackert, 2019). A delivery failure is very costly for e-grocers. The products need to be restored or are spoiled and the delivery needs to be rescheduled. Traditionally, time slots can be differentiated in terms of window length, time of day and price. The traditional incentives are studied widely and also show some drawbacks. First of all, price incentives have a negative impact on the profits (Agatz et al., 2021). Since the last-mile process is the most critical process in the chain, it is important to focus on profit margins. Next to that, if customers are aware that lower prices can occur at a later stage, it is possible that customers delay their purchases which has a negative effect on the sales and planning forecast (Zhang et al., 2020). Thirdly, especially dynamic pricing can be perceived as not fair by the consumers, which can have a negative effect on the trust level of the company (Agatz et al., 2021). Another option is to focus on dynamic slotting, where the number of time slots in a given geographic area is limited in order to cluster customers and reduce the distance or time per order (Agatz et al., 2013). The disadvantage of this approach is that the service proposition becomes less attractive since there are fewer slot options for the customer (Agatz et al., 2021).

In general, there are four classifications of demand management concepts, which are shown in Table 5. The time slot allocation and time slot pricing can be performed static or dynamic (Agatz et al., 2013). *Static* approaches determine the possible options once before the booking horizon starts based on characteristics such as the delivery location and time of day (Mackert, 2019). During the booking horizon, static decisions are not updated. *Dynamic* approaches have as a goal to serve options individually for every customer request that occurs during the booking horizon. Every time window is recalculated in terms of price and window options to customers when a new order is placed. High demand for certain windows, travel time uncertainty and short time windows make this task more complicated (Hays et al., 2005). Dynamic pricing creates more flexibility for the delivery company, but customers can perceive this as unreasonable or unfair. Especially when consumer loyalty plays an important role. An option to include a form of dynamic pricing is to work with discounts or shopping credits next to static pricing. Furthermore, *time slot allocation* focuses on deciding which time slots should be offered to the customers, while *time slot pricing* determines which fee to allocate to which time slot and for which customer, in a dynamic environment (Mackert, 2019).

	Time slot allocation	**Time slot pricing**
Static	Differentiated slotting	Differentiated pricing
Dynamic	Dynamic slotting	Dynamic pricing

Table 5: Classification of demand management concepts (Agatz et al., 2013)

Other, less used, incentives can be rewards such as shopping points or the use of green labels, indicating if the time slot is more sustainable than the other options (Yang et al., 2016). The latter is also defined as environmental benefits. Some time slots are indicated with a green label, which can mean that the corresponding slot is associated with lower fuel consumption and emissions (Agatz et al., 2021). Another definition is that only electric vehicles are used. Using green labels triggers the intrinsic motivation of the customers, while price incentives focus on more extrinsic motivation. Since no impact is made on the thin profit margins, using green labels can be very beneficial for online retailers. Agatz et al. (2021) concluded via a stated choice study on consumer behavior that green labels are an effective tool to steer customers towards a certain delivery option and that it is even more effective on people who are more eco-conscious.

The classification of demand management concepts, as shown in Table 5, can be expanded with the concept of green labeling since it is not part of either time slot allocation or pricing, which is visualized in Table 6. In recent years, companies have increased their offer with regard to sustainable products, giving more insights with regard to polluting emissions and sustainability as a whole is nowadays relevant at every company. Study shows that people are willing to pay more for green electricity and for organic, sustainable food (Nomura and Akai, 2004; Rana and Paul, 2017). Next to that, listing environmental information can have a higher effect on conservation behavior than only financial information (Agatz et al., 2021). It can be concluded that providing green information, on either products or services, has a significant positive effect on the routing performance.

	Time slot allocation	**Time slot pricing**	***Green label***
Static	Differentiated slotting	Differentiated pricing	*Differentiated green label*
Dynamic	Dynamic slotting	Dynamic pricing	*Dynamic green label*

Table 6: Extended classification of demand management concepts (Agatz et al., 2013)

4.4 Time window offering methods at Crisp

The mission of Crisp is to make food of higher quality accessible to a wide audience. The focus is entirely on fresh food products, preferably from local suppliers. The target group is people with busy lives. A minimum order of €50 is required. Crisp is founded in 2018 and is only-app based and available in Dutch. The headquarter and warehouse is located in Amsterdam with hubs in Delft, Utrecht and Eindhoven. In the spring of 2022, they expanded to Belgium, with a warehouse located in Bornem.

The last-mile process of Crisp matches with the process visualized in Figure 7. Only the process in the Netherlands is described. Two times per day, a cut-off moment takes place, at 13:00 and 22:00. Placing an order before 13:00, delivery from the next day in the morning is possible. Placing an order before 22:00, delivery the next day from afternoon is possible. After the cut-off, the in-house developed system makes calculations, where staffing and configuration are done and the information about the orders, such as locations, order size and time windows is uploaded to the route planner. The route planner will plan the routes with the objective to minimize the driving time. When the planning is done, customers receive information about the delivery, such as a track-and-trace and the deliverers receive the route planning. The routes either start from the warehouse in Amsterdam or from one of the hubs. In the latter case, a truck will transport the orders to the hubs in the morning, meaning hubs can only deliver in the afternoon and evening. About 50% of the fleet is electric.

With regard to time windows, Crisp offers time windows with four, two- and one-hour lengths, as shown in Figure 8. The delivery prices are €2.95, €5.95 and €7.95, respectively. When a customer has an order value that is higher than €75, the prices become €2.95 lower, making the four-hour time slot free. Twenty time slots per day are offered.

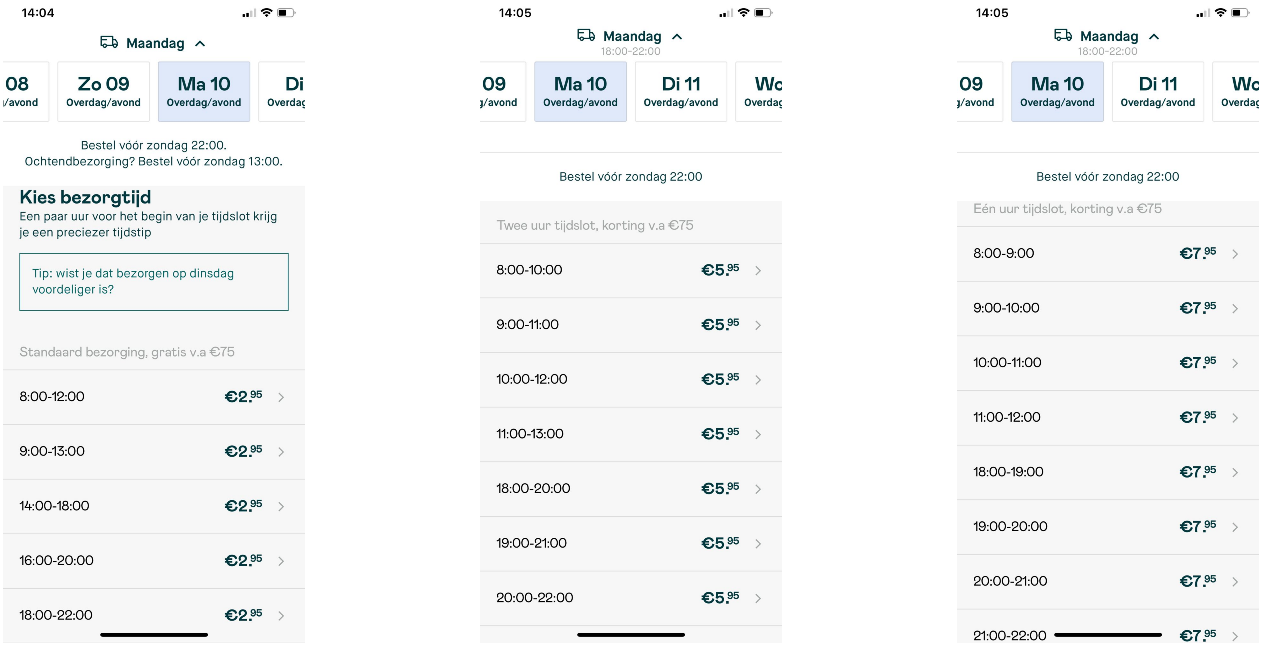


Figure 8: The Crisp app showing the different time slots with corresponding prices

4.5 Time window offering methods at other e-grocers

In Figure 9, the time window offering in the app of Albert Heijn is shown. What is remarkable is that at the top, the sustainable time slots are shown, visualized with a green leaf and the text ‘We are already in the neighborhood’. The sustainable options can differ per day. The width and price of the slots are different for every option every day. Also, full time slots are shown. The offering per day of the week is mainly the same, but the sustainable options can differ. In rare cases, the prices also differ. The time window offering of Albert Heijn can be classified as dynamic; prices per time window can differ per week.

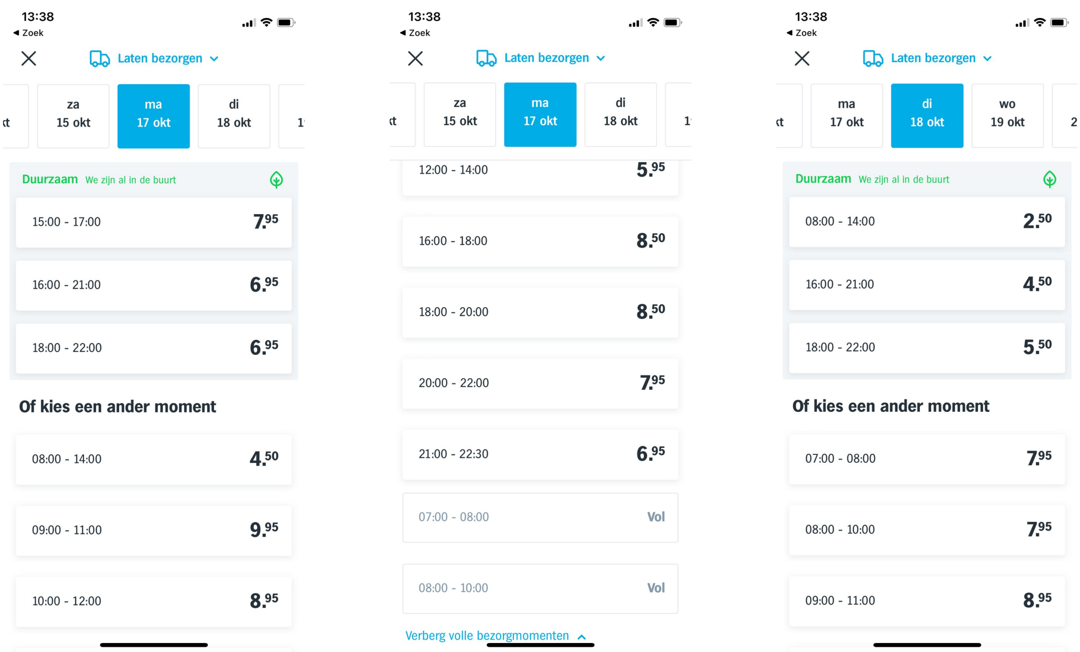


Figure 9: The Albert Heijn app showing the different time slots with corresponding prices

Jumbo has its time slots all visible on one page, making the app more convenient. The offerings and prices differ per day of the week, making the time slot offering of Jumbo dynamic. The different apps were checked every day for over two weeks in November 2022, and the first possible delivery option was four days later at Jumbo. On the 1st of December 2022, the first possibility for delivery is six days later. This makes their time slot offering not user-friendly.

Picnic uses a different strategy than Albert Heijn and Jumbo. Every specific day of the week, the same two options are available. For example, every Thursday, it is possible to get the orders delivered between 08:00 and 09:00 or 17:30 and 18:30, for the delivery location that is inserted. There are no delivery costs. In this way, Picnic is able to cluster its clients very efficiently. They have a minimum order value of €35, making smaller deliveries possible.

As discussed in subsection 4.4, Crisp offers every day 20 slots. This is way higher compared to the other e-grocers, which is shown in Table 7. Albert Heijn offers on average 13 slots per day, with no delivery option on Sunday. Jumbo has on average seven offerings per day and Picnic two. Side note is that this holds for delivery in Amsterdam area.

	Crisp	**AH**	**Jumbo**	**Picnic**
Monday	20	13	7	2
Tuesday	20	13	7	2
Wednesday	20	12.5	6	2
Thursday	20	13	8	2
Friday	20	13	9	2
Saturday	20	13	7	2
Sunday	20	0	4	1

Table 7: Amount of time windows offering per day

In order to compare the delivery costs between the different e-grocers, the average costs per delivery hour are calculated. In order to do so, the price of a time slot is multiplied with the time slot width. For one specific day of the week, the average is taken which is shown in Table 8. The analysis per day of the week are shown in ??. Crisp is the only e-grocer that has prices depending on the order value. The prices differ if the order value is between €50, which is their minimum order value, and €75 or higher than €75. Albert Heijn and Jumbo also have a minimum order value of €50, while Picnic has a minimum value of €35.

	Crisp	AH	Jumbo	Picnic
Monday	€1.93 - €4.90	€6.19	€6.05	€0.00
Tuesday	€1.40 - €3.90	€5.39	€4.81	€0.00
Wednesday	€1.93 - €4.90	€5.70	€3.67	€0.00
Thursday	€1.40 - €3.90	€6.57	€4.88	€0.00
Friday	€1.93 - €4.90	€8.07	€7.69	€0.00
Saturday	€2.15 - €5.04	€5.50	€5.65	€0.00
Sunday	€2.15 - €5.04		€5.47	€0.00
Average	€1.84 - €4.65	€6.24	€5.46	€0.00

Table 8: Average price per hour per day

4.5.1 Sustainability incentives

What can be seen at Albert Heijn, the only grocer in the Netherlands that implemented green labeling, is that every day, three time slots are labeled green of which three time slot options are occurring most often, as can be seen in Table 9. One slot covers the morning and a part of the afternoon (08:00 - 14:00), one covers a part of the afternoon and the evening (16:00 - 21:00) and the last one covers the evening (18:00 - 22:00). Albert Heijn implemented this in 2017 with their partner Ortec (Ortec, 2022). What is interesting to see is that the first two mentioned time slots are covering a part of the afternoon, so when customers opt for these time slots, the delivery can take place in two parts of the day. The afternoon is also seen as the less popular part of the day, so more stops can be pushed towards this day part in order to spread the stops more over the day. Two time slots cover the evening, which is also explainable since the evening can be seen as the most popular part of the day.

The time slot offering at Albert Heijn can be classified as dynamic. Also, the green labeling is dynamic. Sometimes the labeling is not on the three most occurring time slots but on for example 08:00 - 10:00 and 18:00 - 20:00. This is seen when looking more than one week before the potential delivery date. One explanation could be that Albert Heijn wants to cluster that day more toward the morning or evening since that time slot is then used next to another, longer, morning or evening slot.

Start slot	End slot	Width slot	% Occurance
08:00	10:00	2:00	2.9%
08:00	14:00	6:00	27.7%
16:00	21:00	5:00	33.3%
18:00	20:00	2:00	5.6%
18:00	22:00	4:00	30.5%

Table 9: Green labeled time slots at Albert Heijn

Another interesting company that is offering green labeling of time windows recently is PostNL. Here, the definition of sustainable is twofold. It can mean that the parcel will be delivered with an electric vehicle or a vehicle that uses sustainable fuel, or it means that PostNL can make their routing more efficient which has as a consequence that there will be fewer emissions due to fewer driven kilometers. Therefore, pick-up points are, according to PostNL, per definition sustainable. On their one-pager is stated that 66% of Dutch citizens find sustainable delivery important.

One big e-grocer in England is Ocado. Ocado can be seen as an e-grocer but also a technology company. They sell their grocery fulfillment technology to global retailers all over the world. Ocado also used green labeling and is using this method for over five years. Their definition of sustainability is that they can make their routes more efficient based on orders that are already placed for that day (van den Hooven, 2018). Their strategy is interesting since one-hour time slots are offered from 05:30 until 00:00. They offer a 25% discount on your first order and unlimited free deliveries for the first three months.

Sub question 2: What are the different options with regard to time window offering and demand management?

With regard to planning, it is important that the customer is at home during the delivery since a delivery failure is expensive due to the nature of the goods. Therefore, demand management within the e-grocer sector requires attended home delivery (AHD). The standard classification of demand management concepts is subdivided by pricing or time slot allocation incentives, which can occur in a static or dynamic environment. In a static environment, the possible time slot options are determined before the booking horizon starts based on characteristics such as the delivery location and time of day. A dynamic approach has the goal to serve time offering options individually for every customer request that occurs during the booking horizon. A relatively new form of incentive is green labeling, meaning that the labeled time slot will contribute to lower fuel consumption and emissions.

Besides a static or dynamic approach and used incentives, e-grocers can also differentiate their time window offering in terms of the number of time slots. Compared to other e-grocers, Crisp offers a wide variety of time slots and has, besides Picnic, the lowest delivery costs. The time window offering in terms of window length and price occurs dynamically at Albert Heijn and Jumbo. Crisp and Picnic deals with a static strategy, meaning that for a certain delivery location, the offering per day of the week is the same.

Albert Heijn is the only e-grocer in the Netherlands that makes use of green-labeled time slots. The green labeling occurs dynamically, but the greatest share of time slots with a green label are the time slots with a width of at least four hours. Green labeling is also seen at for example PostNL and e-grocers in other countries. A summary of the complete analysis is shown below in Table 10.

E-grocer	**# of time windows**	**Length of time window**	**Costs**	**Green incentive**	**Static/ dynamic**
Crisp	20	1 - 4 h	€0 - €7.95	No	Static
AH	13	1 - 6 h	€3.5 - €9.95	Yes	Dynamic
Jumbo	7	1 - 6 h	€3.5 - €7.95	No	Dynamic
Picnic	2	1 h	€0	No	Static

Table 10: Summary of the competitor analysis

5 Modeling

In this section, the requirements of a choice model specific to this research will be determined. Hereafter, it will be analyzed which choice model suits the requirements best. Next, the preferred choice model will be estimated with the data of Crisp. The results will be served as input for the choice simulations and route optimizations in order to determine the effect of the green time window offering on the routing performance. The following sub-question will be answered:

Sub question 3: How can a choice model be estimated in order to investigate the impact of green labeling on slot choice?

The process from system description to results is shown in Figure 10. This section focuses on choice modeling. In order to estimate a choice model, order data is needed and, in this case, the β estimates on green labeling from Agatz et al. (2021) will be used. The output of the choice model are β estimates on certain time window characteristics which will be the input, together with the order data set for the choice simulation model. With the output from this model, the effect of nudging on the orders can be adjusted in the order data in the test database of an e-grocer and with this, new route simulations can be performed.

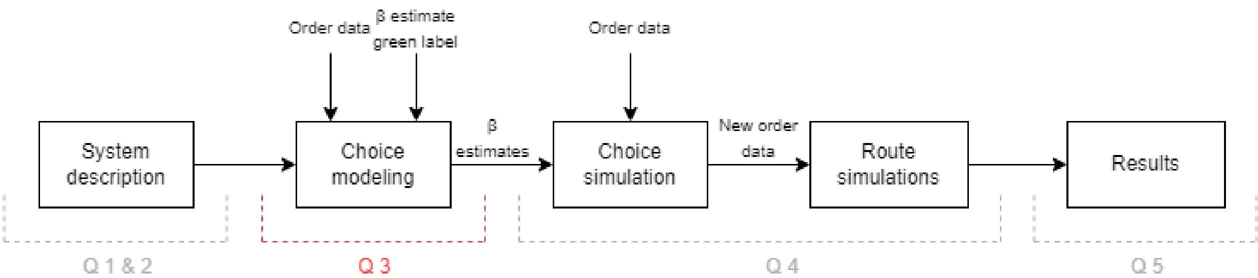


Figure 10: Research framework - Q 3

5.1 Requirement analysis for choice models

In order to choose a suitable choice model, it is helpful to set up several requirements, which can be used to determine which choice model suits best. First of all, it is important that the choice model is able to include *multiple attributes*. So for example, an estimation should not only be based on costs but also on other relevant features. Besides, *computational efficiency* is important. The model must be able to work with large data sets and to optimize in a short period of time, especially when the routing model works within a dynamic environment. The choice model must also be *easy to implement* since the order system and the planning system will use the output of the choice model. The outcomes of how the different choice models score on the requirements are shown in Table 11.

Choice model	**Multiple attributes**	**Computational efficiency**	**Ease of implementation**
Fixed probability	- / +	+	+
Multinomial logit	++	+	+
General attraction	++	-	-

Table 11: Choice model requirements

The explanation of the different choice models can be found in the literature study in section 2. The multinomial logit (MNL) model is due to its flexibility suitable for this research problem. Multiple attributes can be used with different constraints. The resulting coefficients are easy to interpret since it shows the marginal utility of each attribute. An MNL model is also very computationally efficient and easy to implement since it is a linear model. This means that the relationships between the variables are assumed to be linear which makes implementation in a statistical software package more convenient.

5.2 MNL model definition

This subsection will explain how the MNL model is defined and set up. Furthermore, it will explain how the data of the case study is used in order to estimate the attribute values of the MNL model and what its implications are.

According to Agatz et al. (2021), the utility of slot s is $U_s = V_s + \epsilon$, where V_s is equal to $\beta_{green} \cdot X_{s,green}$ in this case. β_{green} is the beta estimates for each label *green* and binary variable $X_{s,green}$ is equal to 1 if slot s has label *green* and 0 otherwise. ϵ stands for the error term which accounts for the unobserved factors related to the time slot preferences. ϵ follows a Gumbel distribution. A β_{green} of 2.8600 will be used from the study of Agatz et al. (2021). This β estimate can be used since the characteristics of the choice model used in that study and the choice model used in this research are very similar.

Objective
The goal of a discrete choice model, specifically an MNL model, is to predict the probabilities of different, categorical options. The model can estimate which option has the highest probability to be chosen, without having to perform experiments in real life. It is multinomial since there are more than two possible discrete outcomes and the dependent variables are nominal.

Assumptions
The MNL model has in general some assumptions. First of all, the dependent variable cannot be completely predicted from the independent variable. Besides, it rests upon the assumption of independence of irrelevant alternatives, meaning that the chance of preferring one option over another option does not depend on other irrelevant alternatives. Another assumption is that the relationship between the independent variables is linear. Furthermore, the multinomial logit model makes the assumption that all category probabilities add up to one and that each category's probability is a function of a linear combination of the independent variables. The independent variables with the highest odds of being observed by the sample's categories are estimated by the coefficients of the model.

Specific to this situation, there are also some assumptions. First of all, it is assumed that β_{green} has a value of 2.8600. In practice, this value can be higher or lower for this specific situation. This cannot be tested due to time constraints. Another assumption is that orders will never arrive simultaneously.

Inputs
Important input that is needed in order to estimate the model is a data set with sufficient orders. Information on certain characteristics, in this case on time window width, costs and starting time is needed.

Function
In order to estimate the needed beta estimates, a script in Apollo, a package of R, will be run. The

complete code can be found in **??** with its output. Different utility functions are written that use the data from the order data set.

Output
The output of the choice model will be a beta estimate for the width, costs and starting times of a time window. Also, the standard error, t-test and p-test values will be given.

5.3 MNL model estimation

In order to fit the MNL to the data of Crisp, more $\beta's$ are estimated. Orders from Crisp are used, resulting in 324 815 orders. When customers opt for a time window, they can choose between different starting times, end times and therefore different window lengths. In this MNL model, the influence of the width of the time window, the starting times on the time slot choice and the costs of the slot are included. This results in the following utility function:

$$U_s = V_s + \epsilon \tag{1}$$

With

$$V_s = \beta_{\text{width}} \cdot \mathrm{X}_{s,\text{ width}} + \beta_{\text{start}} \cdot \mathrm{X}_{s,\text{ start}} + \beta_{\text{costs}} \cdot \mathrm{X}_{s,\text{ costs}} \tag{2}$$

Resulting in:

$$U_s = \beta_{\text{width}} \cdot \mathrm{X}_{s,\text{ width}} + \beta_{\text{start}} \cdot \mathrm{X}_{s,\text{ start}} + \beta_{\text{costs}} \cdot \mathrm{X}_{s,\text{ costs}} + \epsilon \tag{3}$$

In general, Crisp has twenty time windows as shown in Figure 11. First, the different beta's are estimated for the morning orders, followed by the afternoon and evening orders together and lastly from all the orders combined, since some customers will switch to another day part when green labeling will be included. Treating morning orders and afternoon plus evening orders separatelyuy is done since Crisp generates two route planning per day. In this way, it is easier to compare the routing performance.

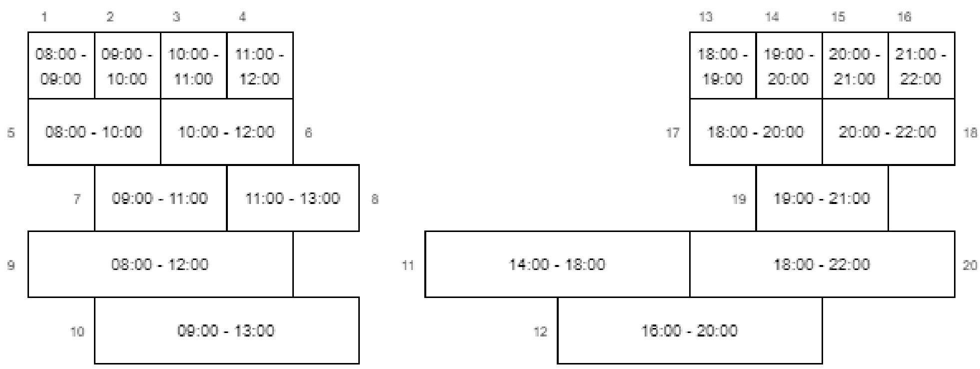


Figure 11: Time slots at Crisp visualised

5.3.1 MNL model morning orders

When preparing the data set, each order id is linked to one of the time slots from 1 until 10, resembling Figure 11. So time slot number 1 is the time slot from 08:00 until 09:00, and therefore has its width equal to one hour ($Width\_1 = 1$), starting time equal to 08:00 ($Start\_1 = 8$) and the costs are equal to €4.95 ($Costs\_1 = 4.95$). This is shown in Table 12. This data set is used as input for an MNL model in Apollo. For the morning delivery, time slots 1 until 10 are included, as shown in Figure 11.

Id	Choice	Width_1	Start_1	Costs_1	...	Width_10	Start_10	Costs_10
1	9	1	8	4.95	...	4	9	0
...	...	...	...	...	...	...	...	...
118 306	10	1	8	4.95	...	4	9	0

Table 12: Data set with morning orders in 2022

The β_{width} has a value of 0.9897, meaning that increasing the window length by one hour will increase the utility on average with 0.9897. This is related to the costs since a wider time window comes with lower costs. β_{start} has a value of -0.3689, meaning that a time slot that starts one hour later will decrease the utility by 0.3689 on average. An explanation can be that customers have a preference for early time slots in the morning. β_{costs} has a value of -0.2576, meaning that an increase of one euro for the delivery costs will lead to a decrease of 0.2576 for the utility of a customer on average. The β_{green} is from the study of Agatz et al. (2021). Therefore, its standard error and t-test values are unknown. A summary of all values is shown in Table 13.

Attribute	Value	Std. error	t-test	p-value	Min	Max
β_{width}	0.9897	0.076039	13.016	0.000	0	0.9897
β_{start}	-0.3689	0.005037	-73.249	0.000	-0.3689	0
β_{costs}	-0.2576	0.049392	-5.215	0.000	-0.2576	0
β_{green}	2.8600			<0.001		

Table 13: Results of the MNL model and characteristics of the attributes for morning deliveries

Now the values of the different attributes are known, the probabilities that a certain time slot is chosen can be calculated with the following formula:

$$P_s = \frac{\exp V_s}{\sum_{j=1}^{j} \exp V_j} \tag{4}$$

Resulting in the following MNL choice probabilities shown in Table 17. This means that, for example, the choice model assigns 36.6% to time window 10, which is the time window from 09:00 to 13:00.

Time slot	1	2	3	4	5	6	7	8	9	10
P(Y=s)	0.8%	0.5%	0.4%	0.3%	3.4%	1.6%	2.4%	1.1%	52.9%	36.6%

Table 14: MNL choice probabilities of morning orders

Since ϵ follows a Gumbel distribution, the error term can be calculated in the following way: $-LN(-LN(X))$ where X is a randomized number between 0 and 1. Due to this randomization in the error term, the estimated choices can deviate from the choice probabilities as shown in Table 17.

5.3.2 MNL model afternoon and evening orders

When preparing the data set, each order id linked to one of the time slots from 11 until 20, resembling Figure 11. So time slot number 11 is the time slot from 14:00 until 18:00, and therefore has its width equal to four hours ($Width\_1 = 4$), starting time equal to 14:00 ($Start\_1 = 14$) and the costs are equal to €0.- ($Costs\_1 = 0$). This is shown in Table 15. For the static simulation, only the afternoon and evening delivery is included, which are time slots 11 until 20 in Figure 11.

Id	Choice	Width_11	Start_11	Costs_11	...	Width_20	Start_20	Costs_20
1	19	4	14	0	...	4	18	0
...	...	...	...	...	...	...	...	...
206 509	20	4	14	0	...	4	18	0

Table 15: Data set with afternoon and evening orders in 2022

The β_{width} has a value of 1.9713, meaning that increasing the window length by one hour will increase the utility on average with 1.9713. This is related to the costs since a wider time window comes with lower costs. β_{start} has a value of 0.1080, meaning that a time slot that starts one hour later will increase the utility by 0.1080 on average. An explanation can be that customers have a preference for later time slots during the afternoon and evening. β_{costs} has a value of 0.1979, meaning that an increase of one euro for the delivery costs will lead to an increase of 0.1979 for the utility of a customer on average. This is counterintuitive but since the costs are related to the width of the time windows, it is reasonable. A summary of all values is shown in Table 16.

Attribute	Value	Std. error	t-test	p-value	Min	Max
β_{width}	1.9713	0.073270	26.904	0.000	0	1.9713
β_{start}	0.1080	0.001392	77.629	0.000	0	0.1080
β_{costs}	0.1979	0.047268	4.188	0.000	0	0.1979
β_{green}	2.8600			<0.001		

Table 16: Results of the MNL model and characteristics of the attributes for afternoon and evening deliveries

Now the values of the different attributes are known, the probabilities that a certain time slot is chosen can be calculated, resulting in the following MNL choice probabilities:

11	12	13	14	15	16	17	18	19	20
29.4%	31.3%	0.3%	0.3%	0.3%	0.4%	1.5%	1.6%	1.5%	33.3%

Table 17: MNL choice probabilities of afternoon and evening orders

5.3.3 MNL model whole day

When preparing the data set, each order id is linked to one of the time slots from 1 until 20, resembling Figure 11. So time slot number 1 is the time slot from 08:00 until 09:00, and therefore has its width equal to one hour ($Width\_1 = 1$), starting time equal to 08:00 ($Start\_1 = 8$) and the costs are equal to €4.95 ($Costs\_1 = 4.95$). This is shown in Table 18. For the static simulation, all the time windows are included, which are time slots 1 until 20 in Figure 11.

Id	Choice	Width_1	Start_1	Costs_1	...	Width_20	Start_20	Costs_20
1	9	1	8	4.95	...	4	18	0
...	...	...	...	...	...	...	...	...
324 815	20	1	8	4.95	...	4	18	0

Table 18: Data set with all orders in 2022

The β_{width} has a value of 1.6112, meaning that increasing the window length by one hour will increase the utility on average with 1.6112. This is related to the costs since a wider time window comes with lower costs. β_{start} has a value of 0.0313, meaning that a time slot that starts one hour later will increase the utility with 0.0313 on average. β_{costs} has a value of 0.0415, meaning that an increase of one euro for the delivery costs will lead to an increase of 0.0415 for the utility of a customer on average. The β_{green} is from the study of Agatz et al. (2021). Therefore, its standard error and t-test values are unknown. A summary of all values is shown in Table 19.

Attribute	Value	Std. error	t-test	p-value	Min	Max
β_{width}	1.61123	0.05248	30.700	0.000	0	1.61123
β_{start}	0.03134	0.00045	70.262	0.000	0	0.03134
β_{costs}	0.04146	0.03402	1.218	0.1115	0	0.04146
β_{green}	2.8600			<0.001		

Table 19: Results of the MNL model and characteristics of the attributes for deliveries during the whole day

Now the values of the different attributes are known, the probabilities that a certain time slot is chosen can be calculated, resulting in the following MNL choice probabilities:

Time slot	1	2	3	4	5	6	7	8	9	10
P(Y=s)	0.2%	0.2%	0.2%	0.2%	0.7%	0.8%	0.7%	0.8%	15.7%	16.2%

11	12	13	14	15	16	17	18	19	20
19.0%	20.2%	0.2%	0.2%	0.2%	0.2%	1.0%	1.0%	1.0%	21.5%

Table 20: MNL choice probabilities of all orders

5.4 Validation of the MNL model

In order to analyze the performance of the choice model estimation, the number of times that a certain time slot is chosen will be compared with the outcomes of the choice model estimation. An overview is shown in Table 21. It is interesting to see that the differences between the real chosen time slots and the estimated ones become considerably smaller from the time window 5 until 10, which is the first time window with a width of two hours. For the choice model, it is hard to simulate the shortest time window of one hour correctly. This is explainable since only a few customers choose the smallest time windows and according to the MNL model, this is linearly hard to estimate.

Time slot	1	2	3	4	5	6	7	8	9	10
Real	1217	545	264	221	4 333	1 739	2 439	1 608	62 054	43 886
Estimated	3	1	425	310	4 152	1 960	2 820	1 339	63 286	43 755
Relative difference	-99.8%	-99.8%	61.0%	40.3%	-4.2%	12.7%	15.6%	-16.9%	2.0%	-0.3%

Table 21: Amount of times the time slots are chosen

The next step is to determine how many choices will change when including the green label attribute. In this stage, this is only possible for the static environment. In order to do so $\beta_{green} \cdot X_{s,green}$ will be added to the utility function, resulting in:

$$U_s = \beta_{\text{width}} \cdot \mathrm{X}_{s,\text{ width}} + \beta_{\text{start}} \cdot \mathrm{X}_{s,\text{ start}} + \beta_{\text{costs}} \cdot \mathrm{X}_{s,\text{ costs}} + \beta_{green} \cdot X_{s,green} + \epsilon \tag{5}$$

In Table 22 can be seen that the green-labeled slots, which are slot 9 and 10, increase with 6 044 and 4 298, respectively.

Time slot	1	2	3	4	5	6	7	8	9	10
Real	1217	545	264	221	4 333	1 739	2 439	1 608	62 054	43 886
Estimated	3	1	425	310	4 152	1 960	2 820	1 339	63 286	43 755
Estimated green	0	0	27	15	248	123	174	81	69 330	48 053

Table 22: Amount of times the time slots are chosen including the influence of green labeling

In Table 23 can be seen that the green-labeled slots increase with 5.1% and 3.6%, respectively. This is a total increase of 9.4% when looking at total orders.

Time slot	1	2	3	4	5	6	7	8	9	10
Real	1.0%	0.5%	0.2%	0.2%	3.7%	1.5%	2.1%	1.4%	52.5%	37.1%
Estimated	0.0%	0.0%	0.4%	0.3%	3.5%	1.7%	2.4%	1.1%	53.6%	37.1%
Estimated green	0.0%	0.0%	0.0%	0.0%	0.2%	0.1%	0.1%	0.1%	58.7%	40.7%

Table 23: Amount of times the time slots are chosen relatively including the influence of green labeling

Sub question 3: How can a choice model be estimated in order to investigate the impact of green labeling on slot choice?

Different choice models are defined and scored against different requirements: *multiple attributes*, *computational efficiency* and *easy to implement*. This results in the following outcome and therefore, the MNL model will be used since it is the best fit for this study:

Choice model	Multiple attributes	Computational efficiency	Ease of implementation
Fixed probability	- / +	+	+
Multinomial logit	++	+	+
General attraction	++	-	-

Table 24: Choice model requirements

An MNL model will predict the probabilities of different options. It is multinomial since there are more than two possible discrete outcomes and the dependent variables are nominal. First, different beta estimates are calculated in Apollo with the data of Crisp. These beta estimates are used in the utility function without the green label attribute in order to estimate the choice probabilities of the MNL model. This is done for the morning deliveries, the afternoon and evening deliveries together and for all the orders on the same day. A summary is shown in Table 25. Thereafter, the beta estimate of applying a green label is included. This has resulted in larger time slots being chosen more often. In total, there is an increase of 9.4% in large time windows.

Attribute	Morning	Afternoon & evening	All day
β_{width}	0.9897	1.9713	1.6112
β_{start}	-0.3689	0.1080	0.0313
β_{costs}	-0.2576	0.1979	0.0415

Table 25: Summary of the output of the choice model

6 Simulation

In the previous section, multiple choice models are defined and scored against different requirements. The best fitting model, the MNL model, is estimated with real data from an e-grocer. The output of the MNL model will be used in this section for the choice simulation in a static and dynamic environment. For the static environment, less detailed requirements are needed since larger time windows lead to a more efficient and thus 'greener' solution, which will be specified later on. This chapter ends with an experimental plan in subsection 6.3 where the settings and other simulation characteristics are explained. In this section, the following sub-question will be answered:

Sub question 4: How can a simulation study be set up in order to study the effect of green label time window offering on the routing performance?

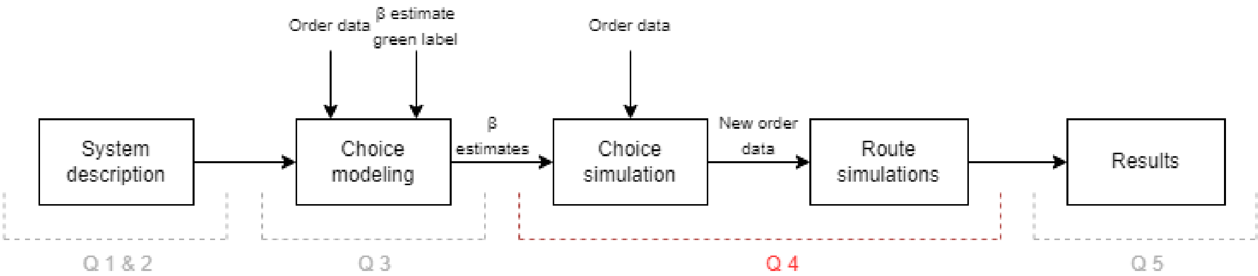


Figure 12: Research framework - Q 4

6.1 Static incentive method

In a static environment, it is beforehand determined which time windows will receive an incentive. Different approaches are possible. One approach is to give all large time windows a green label. What also is possible, is that not all delivery locations will receive the same time windows to choose from. It is possible that for example, certain low-density areas can choose only evening deliveries. This is a way of clustering. In Figure 13 is the process of an arriving customer in a static time slot management (STSM) system shown.

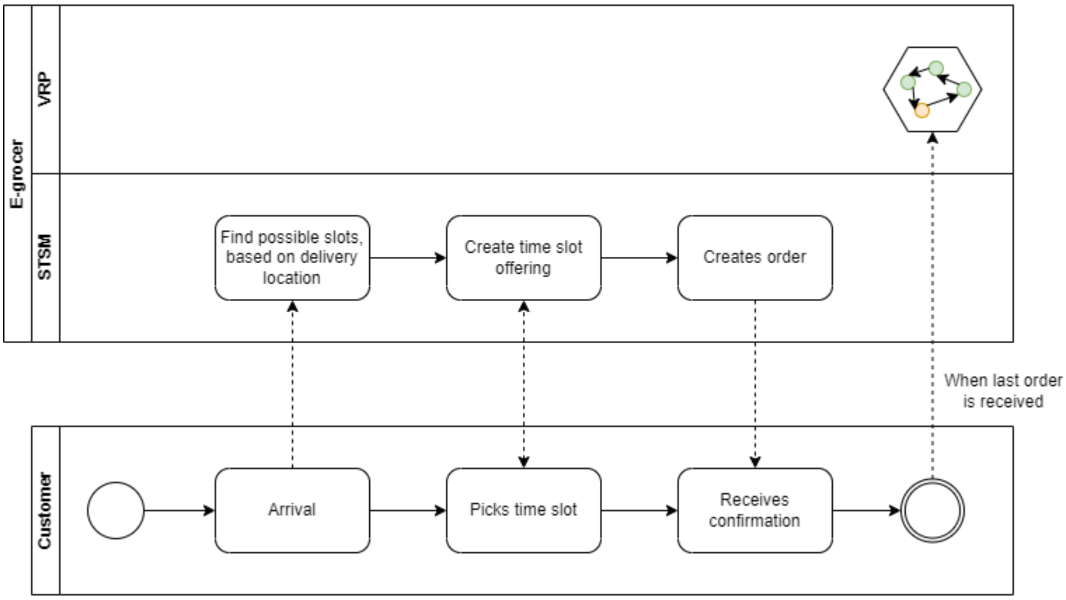


Figure 13: Process of an arriving customer in a STSM system

Objective

The objective of an STSM system with green labeling is to nudge customers towards more preferred time windows. In static modeling, this means that the larger time windows get a green label. The objective is that more customers opt for a green, thus larger, window than the scenario with no green labeling. The hypothesis is that this should increase the routing performance compared to a situation with no green labeling, while it should have a lower effect than dynamic green labeling.

Assumptions

There are multiple assumptions regarding this scenario. First of all, it is assumed that there is enough capacity to serve the customers after the nudging process. Meaning that no extra costs are needed when for example customers switch to another day part. Another assumption is that the behavior of the customer follows the choice model as described in section 5.

Input

In order to simulate the effect of the static green labeling process, the historic orders need to be adjusted according to the outcomes of the MNL model. In order to do so, the following information is needed:

- Historic order data
- Characteristics of time windows
- Output of choice model

Function

Figure 13 shows how the STSM system looks like when implemented. The following steps need to be performed in order to prepare the data set:

1. Filter the MNL model in Excel on the to-be-simulated day

2. Filter on which orders changed due to the nudging
3. Change time windows of those orders in the order system
4. Run a new route planning in the routing software system

Output

The output is a new route planning which can be compared to the route planning with historic data. The output will be scored according to the defined KPIs from subsection 4.2, *costs* and *sustainability*, in order to assess the performance of the route plan.

6.1.1 Static labeling approaches

The approaches shown in Table 26 will be simulated, resulting in two approaches. By analyzing the results per day part and for the whole day, it can be tested what the effect is when customers stay with their original day part and when they are allowed to switch to another. In approach 1, all customers are offered the same time windows and therefore the same time windows with a green label. In approach 2, the green labeled four-hour time windows are replaced by five-hour time windows. This is done in order to analyze the effect of a bigger time window.

Approach	Morning	Choice model
#1	All four hour time windows receive a green label	- Morning - Afternoon + evening - All day
#2	All five hour time windows receive a green label	- Morning - Afternoon + evening - All day

Table 26: Approaches of the static simulations

A visualization of which time windows receive a green label is shown in Figure 14. The four-hour time slots receive a green label, which is visible to the customer at the check-out of the order process. In case of five-hour time windows, time windows 9 and 10 become 08:00 - 13:00, time window 11 changes to 14:00 - 19:00, time window 12 to 16:00 - 21:00 and time window 20 become 17:00 - 22:00.

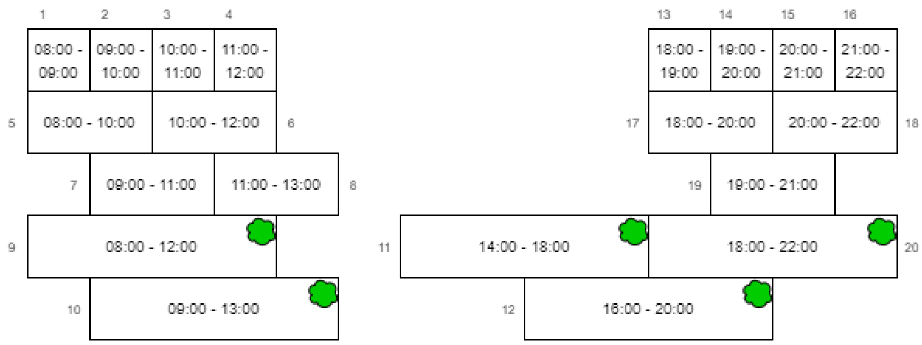


Figure 14: Green label on time windows in a static environment

6.2 Dynamic incentive method

There are multiple ways to implement a dynamic time slot management (DTSM) system. Some of the options are visualized and explained below:

Intermediate VRP optimisation
In this version of the DTSM, a system creates a time slot offering for customers that arrive in the system. So in the case of an e-grocer, this happens mostly when the products are placed in the basket. The time slot offering in a dynamic environment is based on a current delivery schedule which is periodically optimized. This can be after each order or a few times per day, depending on the computational speed (Yang et al., 2016). The customer picks a time slot, whereafter the customer receives an order confirmation and the delivery route is re-optimized. The current delivery schedule is a solution to the vehicle routing problem (VRP). The process is shown in Figure 15.

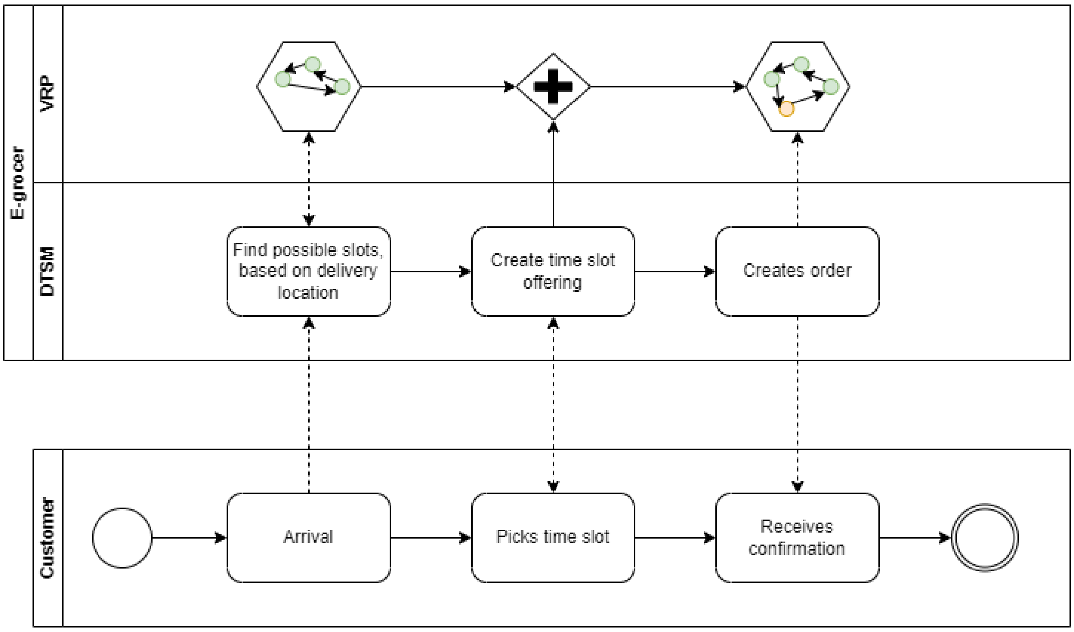


Figure 15: Process of an arriving customer in a DTSM system

VRP optimisation after final cut-off
In this version of the DTSM, a system creates a time slot offering for customers that arrive in the system. The green labeling of the time slot offering in this dynamic environment is based on previously accepted orders. When a new customer arrives and there is already an accepted order from a previous customer in the same area, green labeling will be applied based on that. The definition of an area can be via zip coding or distance between customers (van Mil, 2022). The process is shown in Figure 16.

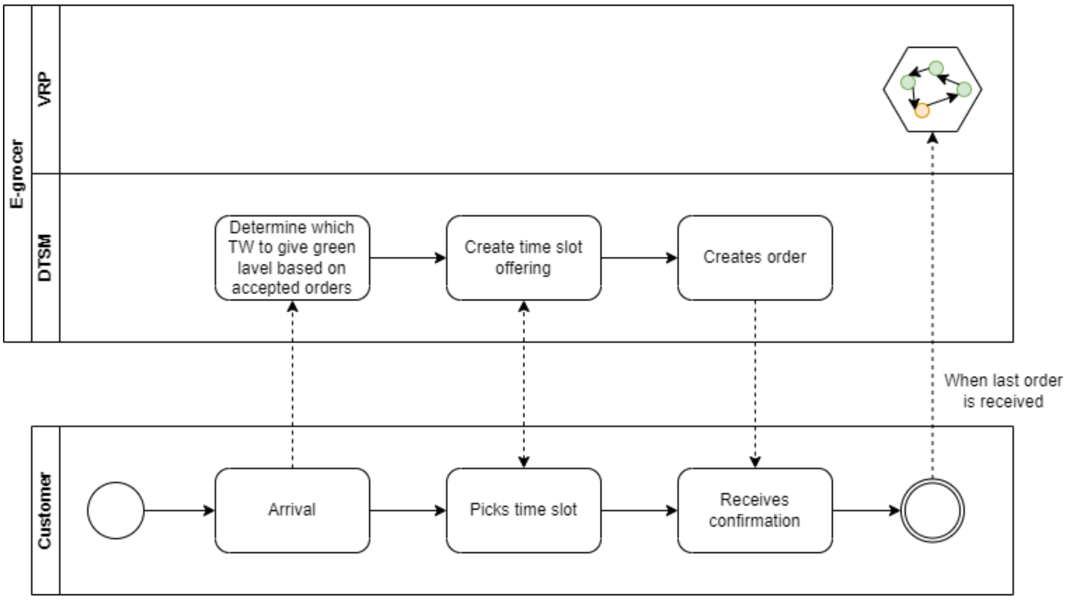


Figure 16: Process of an arriving customer in a DTSM system which optimize after cut-off

Since this research uses commercial routing software for the simulations, the green labeling should be determined on the locations of the already accepted orders and not on by means of intermediate routing optimization after each order. This means that the process visualized in Figure 16 will be followed.

Therefore, green labeling will be done based on the zip code of the accepted orders. Using this method, no distance matrix is used or other ways of determining the distance between two or multiple points. Labeling based on zip code can be done for multiple levels. PC1 level, which means that for example, zip code 1234AB becomes simply 1, results in a too large area and will therefore not be included. This means that the Netherlands will be divided into nine large areas. PC2 level results in around 90 areas, which can be efficient for the labeling process and will therefore be included. PC3, with 900 areas and PC4, with 9000 areas, will result in a too detailed separation of areas and therefore not included.

Objective

The objective of a DTSM system with green labeling is to steer customers towards more preferred time windows. In dynamic modeling, this means that time windows that get a green label depend on already accepted orders. The objective is that customers will be automatically better grouped in terms of their chosen time window with respect to their location. This should improve the routing performance.

Assumptions

There are multiple assumptions regarding this scenario. First of all, it is assumed that there is enough capacity to serve the customers after the nudging process. Meaning that no extra costs are needed when for example customers switch to another day part. Another assumption is that the behavior of the customer follows the choice model as described in section 5. Besides, an assumption is that a green label on a smaller time window is possible and does not negatively affect the routing costs if an earlier customer in that area already placed an order in a smaller time window.

Input
In order to simulate the effect of the dynamic green labeling process, the historic orders need to be adjusted according to the outcomes of the MNL model. In order to do so, the following information is needed:

- Historic order data
- Characteristics of time windows
- Output of the choice model

Function
Figure 15 shows how the DTSM system looks like when implemented. For simulation purposes, the steps are listed below:

1. Filter the MNL model in Excel on the to-be-simulated day
2. Follow one of the flow charts for the different approaches as described in subsubsection 6.2.1. Repeat this for every customer
3. Filter on which orders changed due to the nudging
4. Change time windows of those orders in the order system of the e-grocer
5. Run a new route planning in the routing software system

Output
The output is a new route planning which can be compared to the route planning with historic data. The output will be scored according to the defined KPIs from subsection 4.2, *duration*, *routes/vehicles* and *distance*, in order to assess the performance of the route plan.

6.2.1 Dynamic labeling approaches

In this subsection, two dynamic simulation approaches will be described, together with its flowchart.

Approach 1
There are multiple approaches in order to performing dynamic labeling. The first approach is visualized in Figure 17. A distinction is made between low-density and high-density areas in the Netherlands. The reasoning behind this is that in high-density areas, there will be many delivery vans serving those areas all day, therefore the exact green-labeled time windows matter less and will for computational reasons be treated as a static scenario. A difference is that this approach only places a green label on the morning and afternoon four-hour time window in order to nudge customers more towards the less popular day parts.

For the low-density areas, first will be checked if there is an already accepted order in that zip code with the associated level. If not, the same policy as for the low-density area will hold. This is followed by a distinction in how many day parts the accepted orders are and how many accepted orders are in that day part.

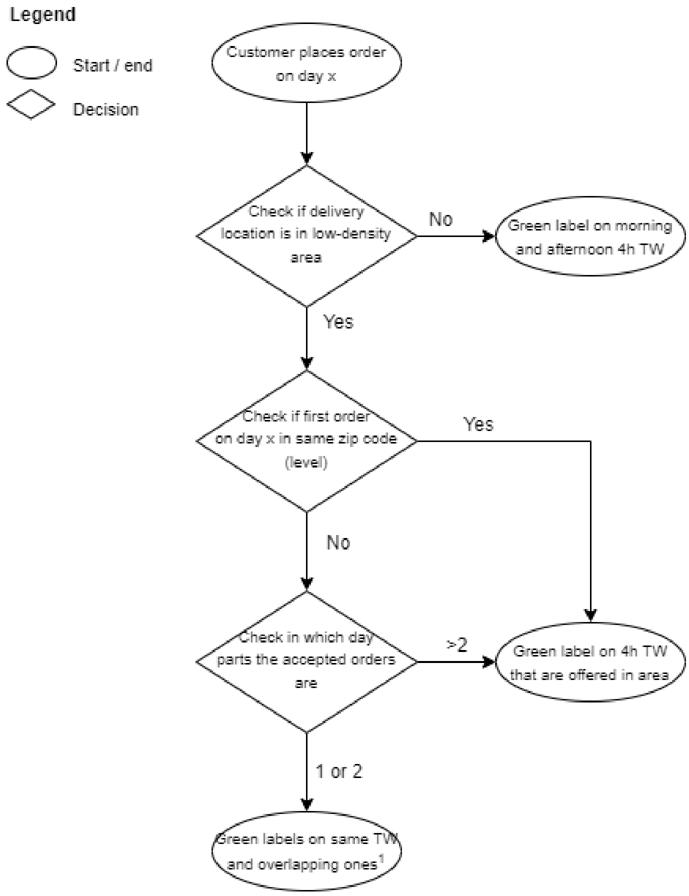


Figure 17: Flow chart for simulation approach 1

Approach 2

The second dynamic approach is visualized in Figure 18. A distinction is made between low-density and high-density areas in the Netherlands. This approach is quite similar to the previous one, except that customers are not nudged more towards the morning or afternoon. High-density areas will get a green label on all four-hour time slots, similar to the static approach, while for the high-density areas, the same policy holds as for the first dynamic approach.

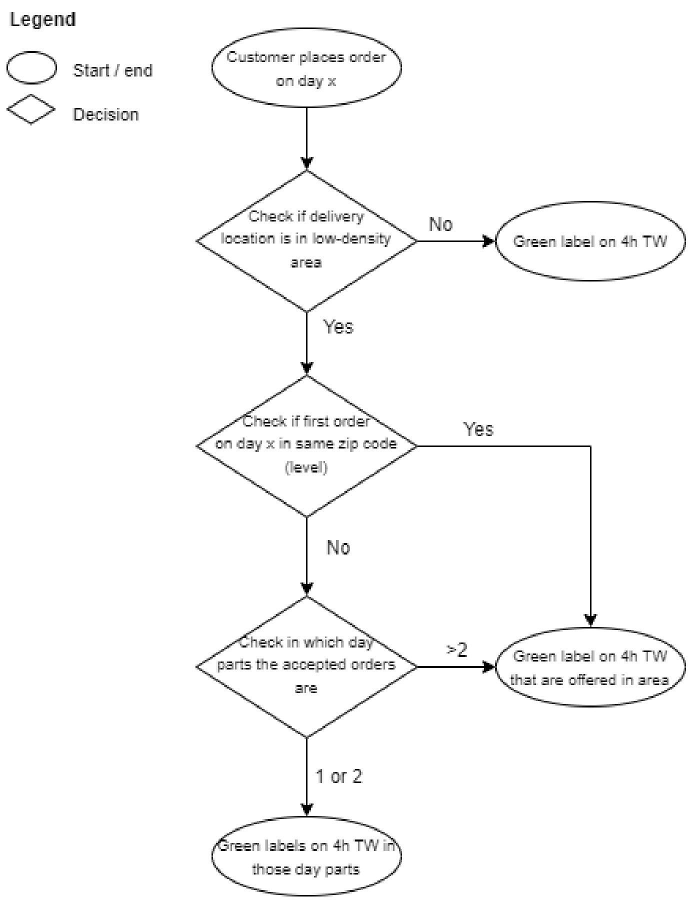


Figure 18: Flow chart for simulation approach 2

6.3 Experimental plan

This subsection will describe the experimental plan; the most important inputs and their definition in order to come to successful simulation experiments. In total, four approaches are designed, each with five days to experiment. This leads to a total of 20 experiments.

First, the grouping parameters will be described, which is the input for the choice simulation. Also, the various planning parameters will be described, explaining the parameters needed for the route simulations. The relations to each other are shown in the experimental set-up below in Figure 19.

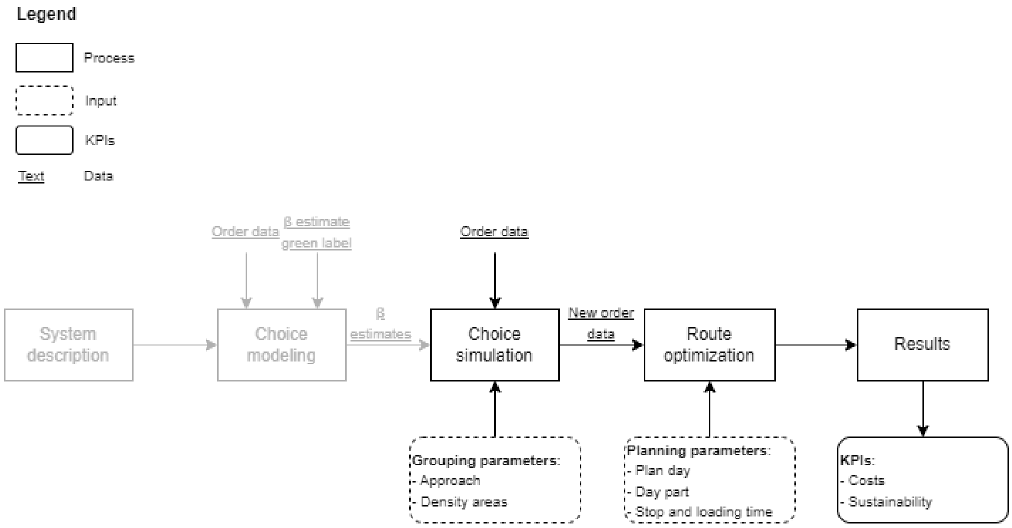


Figure 19: Experimental set-up

Grouping parameters

For every simulation run, the choice simulation needs to know which *approach* is used, which can be either one of the two static or dynamic approaches, as described in subsection 7.1 and subsection 7.2. Also, the *density area* must be specified, the high-density areas are in this study set to the PC2 areas 10, 11, 12, 15, 19, 20, 21, 22, 23, 25, 26, 30, 35 and 37 since these areas have the highest concentration of order according to the data set.

Planning parameters

For the route simulations, the route planner needs to know which *plan day* needs to be simulated. In order to have reliable results, five plan days are simulated for each approach. Furthermore, it must be clear what the *day part* is that has to be simulated. Due to the nature of the data set, it is only possible to simulate only the morning or the afternoon and evening together. Therefore, these two variants can be seen in the results, together with simulations of the whole day which is representative of the results of the morning and afternoon plus evening combined. Lastly, the *stop time* and *loading time* must be inserted. The *stop time* is set to five minutes for every stop, while the *loading time* varies between 15 and 30 minutes.

Sub question 4: How can a simulation study be set up in order to study the effect of green label time window offering on the routing performance?

In this research, the simulation study will be set up in a static and dynamic way. In static time slot management (STSM), the larger time windows receive a green label. Here, two approaches are defined. In the first approach, all the four-hour time windows receive a green label. In the second approach, the four-hour time windows are changed to five-hour time windows. Here, there is accounted for that in for example low density areas, only evening deliveries are performed.

Within dynamic incentive methods, there are two main directions; *intermediate vehicle routing problem (VRP) optimization* and *VRP optimization after final cut-off*. Since this research uses commercial routing software for the simulations, the green labeling should be determined on the locations of the already accepted orders and not on by means of intermediate routing optimization after each order. Therefore, the *VRP optimization after final cut-off* will be used. The green labeling will be done based on the zip codes of the accepted orders; customers receive adapted green labels when their location is within a certain proximity of already accepted orders for that day. There will be two dynamic labeling approaches simulated. The first approach has almost the same rules for high-density areas compared to the static situation. A difference is that this approach only places a green label on the morning and afternoon four-hour time window in order to nudge customers more towards the unpopular day parts For low-density areas, different rules will apply on which time window receives a green label. The second approach is similar, except that customers are not nudged towards the morning or afternoon.

The grouping parameters of the choice simulation are one of the four approaches and a definition of the density areas. The grouping parameters of the route optimization are the plan day, day part and stop- and loading time.

7 Results

In this chapter, the results of various choice simulation approaches are described via route optimization. subsection 7.1 describes the results of the different approaches with regard to the static scenario and in subsection 7.2, the dynamic results are described and analyzed. Each approach is expressed in terms of costs and sustainability, conforming to the KPIs as defined in subsection 4.2.

In this section, the following sub-question will be answered:

Sub question 5: What is the effect of green label time window offering on the routing performance?

7.1 Static green labeling

In order to analyze the effect of static green labeling on the routing performance, multiple days are optimized. In the following tables, 'duration' stands for the total time needed for all the routes, 'routes' shows the number of routes created for that day and 'distance' for the total route kilometers. First, the relative improvement per day of the three factors will be treated per approach. This will be done for the different day parts and for the whole day. This is done since this gives a good impression of the difference in the effect of the multiple approaches. These relative improvements can be converted to monetary- and emission units, resulting in the KPIs at the end of each subsection.

Approach 1 - morning

In Table 27, the results are shown of the morning optimization of the first approach within static green labeling. In this approach, all four-hour time windows receive a green label. On average, the duration decreases, and therefore improves, by 2.9%, the amounts of routes decrease by 2.6% and the distance decreases by 4.2%. These results are compared to the results in the base case, where no adjustments are made. Since the number of orders stays the same, the relative improvements hold for both total improvements and improvements per order.

	Day #1	**Day #2**	**Day #3**	**Day #4**	**Day #5**	**Average**
Duration	-3.7%	-1.7%	-3.7%	-2.5%	-3.0%	-2.9%
Routes	-3.6%	0.0%	-4.4%	-3.0%	-2.2%	-2.6%
Distance	-3.6%	-4.1%	-5.3%	-4.2%	-3.7%	-4.2%

Table 27: Optimization results morning static green labeling - approach 1

Approach 1 - afternoon & evening

In Table 28, the results are shown of the afternoon plus evening optimization of the first approach within static green labeling. On average, the duration improves by 3.1%, the amount of routes decreases 3.4% and the distance decreases by 3.7%.

	Day #1	**Day #2**	**Day #3**	**Day #4**	**Day #5**	**Average**
Duration	-4.1%	-1.9%	-3.8%	-2.9%	-3.0%	-3.1%
Routes	-3.7%	-3.0%	-4.1%	-3.5%	-2.6%	-3.4%
Distance	-3.2%	-3.5%	-4.3%	-4.5%	-3.2%	-3.7%

Table 28: Optimization results afternoon and evening static green labeling - approach 1

Approach 1 - whole day

In Table 29, the results are shown of the optimizations from the whole day of the first approach within static green labeling. This means that the outcomes of the route optimizations are summed. On average, the duration improves by 3.1%, the number of routes decreases also by 3.1% and the distance decreases by 3.9%. The results are also visualized in Figure 20. Only the whole day optimization is visualized in order to compare to all different approaches.

	Day #1	Day #2	Day #3	Day #4	Day #5	Average
Duration	-4.0%	-1.9%	-3.8%	-2.8%	-3.0%	-3.1%
Routes	-3.7%	-2.1%	-4.1%	-3.3%	-2.5%	-3.1%
Distance	-3.3%	-3.7%	-4.6%	-4.4%	-3.4%	-3.9%

Table 29: Optimization results whole day static green labeling - approach 1

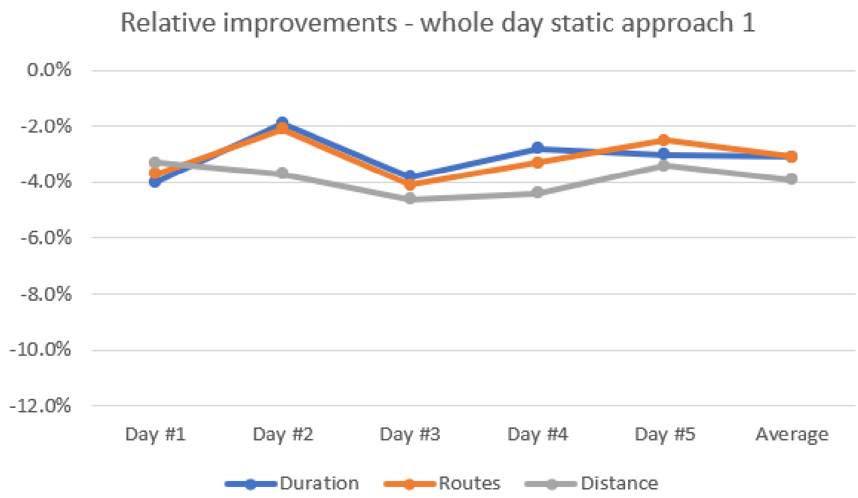


Figure 20: Relative improvements - whole day static approach 1

Approach 2 - morning

In the second approach, the four-hour green labeled time slots are replaced by five-hour time slots as described in section 6. In Table 30, the results are shown of the optimizations from the morning of the second approach within static green labeling. On average, the duration improves by 3.9%, the number of routes decreases by 3.1% and the distance decreases by 5.1%.

	Day #1	Day #2	Day #3	Day #4	Day #5	Average
Duration	-4.7%	-2.5%	-4.8%	-3.7%	-3.9%	-3.9%
Routes	-3.7%	0.0%	-5.0%	-3.9%	-2.9%	-3.1%
Distance	-4.6%	-4.9%	-6.2%	-5.1%	-4.8%	-5.1%

Table 30: Optimization results morning static green labeling - approach 2

Approach 2 - afternoon & evening

In Table 31, the results are shown of the optimizations from the afternoon and evening of the second approach within static green labeling. On average, the duration improves by 3.9%, the amounts of routes decrease by 4.4% and the distance decreases by 4.7%.

	Day #1	Day #2	Day #3	Day #4	Day #5	Average
Duration	-4.0%	-2.7%	-5.0%	-3.8%	-3.8%	-3.9%
Routes	-4.9%	-4.1%	-4.9%	-4.4%	-3.5%	-4.4%
Distance	-3.9%	-4.5%	-5.2%	-5.6%	-4.1%	-4.7%

Table 31: Optimization results afternoon and evening static green labeling - approach 2

Approach 2 - whole day

In Table 32, the results are shown of the optimizations from the whole day of the second approach within static green labeling. This means that the outcomes of the route optimizations are summed. On average, the duration improves by 3.9%, the number of routes decreases by 4.0% and the distance decreases by 4.8%. The results are also visualized in Figure 21.

	Day #1	Day #2	Day #3	Day #4	Day #5	Average
Duration	-4.9%	-2.7%	-4.8%	-3.7%	-3.4%	-3.9%
Routes	-4.8%	-2.6%	-5.1%	-4.1%	-3.3%	-4.0%
Distance	-4.2%	-4.8%	-5.5%	-5.6%	-3.8%	-4.8%

Table 32: Optimization results whole day static green labeling - approach 2

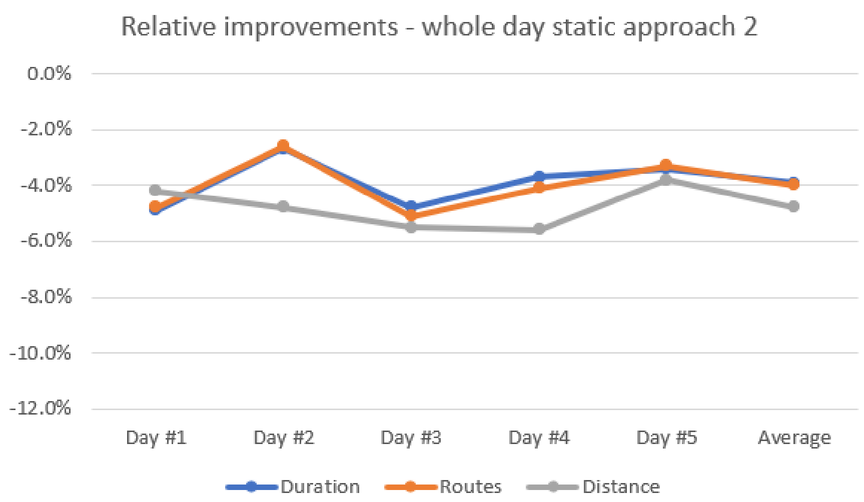


Figure 21: Relative improvements - whole day static approach 2

7.1.1 KPIs

In order to compare both scenarios on costs, *duration*, used *vehicles* and *distance* are changed to monetary improvement, as described in subsection 4.2. With regard to sustainability, only *distance* is formulated in emissions. Only the results from the whole day are included since analyses by day part can give a wrong impression. In some approaches, orders will switch to another day part. In this way, comparisons are more reliable. Table 33 shows the improvement in costs per order. So on average, the costs per order become €0.53 lower in the first static approach.

	Day #1	Day #2	Day #3	Day #4	Day #5	Average
Duration	€0.41	€0.18	€0.38	€0.65	€0.48	€0.41
Routes	€0.08	€0.00	€0.06	€0.15	€0.08	€0.07
Distance	€0.06	€0.04	€0.04	€0.09	€0.07	€0.05
Total	€0.55	€0.22	€0.48	€0.91	€0.48	**€0.53**

Table 33: Impact on costs from optimizations - whole day static approach 1

Table 34 shows the impact of distance on sustainability in terms of emissions. On average, 46.2 g CO_2 less is emitted per order.

	Day #1	Day #2	Day #3	Day #4	Day #5	Average
Distance	33.0 g CO_2	40.1 g CO_2	58.7 g CO_2	50.2 g CO_2	49.1 g CO_2	**46.2** g CO_2

Table 34: Impact on sustainability from optimizations - whole day static approach 1

The KPIs are also visualized in Figure 22, where the left axis indicated the monetary improvements per order and the right axis shows the improvement per order in terms of grams CO_2.

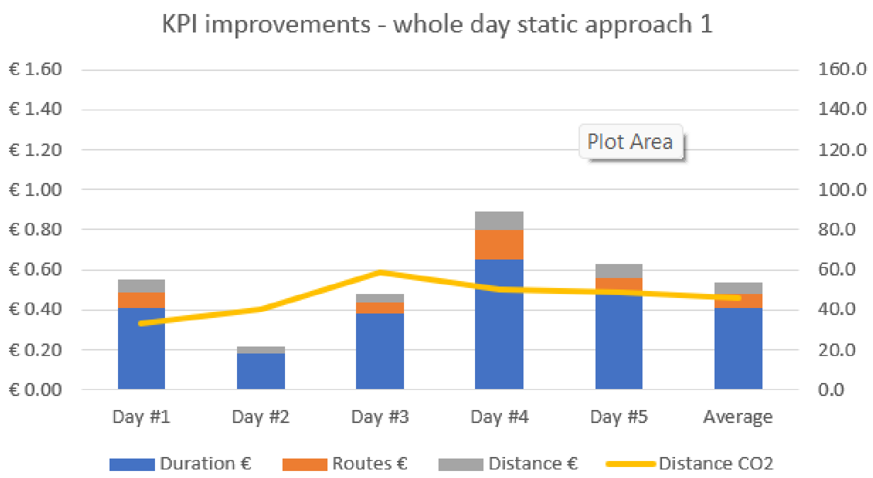


Figure 22: KPI improvements - whole day static approach 1

Table 35 shows the improvement in costs per order. So on average, the costs per order become €0.63 lower in the second static approach.

	Day #1	Day #2	Day #3	Day #4	Day #5	Average
Duration	€0.45	€0.20	€0.44	€0.70	€0.53	€0.46
Routes	€0.10	€0.07	€0.08	€0.17	€0.09	€0.10
Distance	€0.07	€0.05	€0.05	€0.10	€0.08	€0.07
Total	€0.62	€0.32	€0.57	€0.97	€0.70	**€0.63**

Table 35: Impact on costs from optimization - whole day static approach 2

Table 36 shows the impact of distance on sustainability in terms of emissions. On average, 55.4 g CO_2 less is emitted per order. The KPIs for the second approach are also visualized in Figure 23.

	Day #1	Day #2	Day #3	Day #4	Day #5	Average
Distance	37.4 g CO_2	46.3 g CO_2	79.6 g CO_2	61.1 g CO_2	52.8 g CO_2	**55.4** g CO_2

Table 36: Impact on sustainability from optimizations - whole day static approach 2

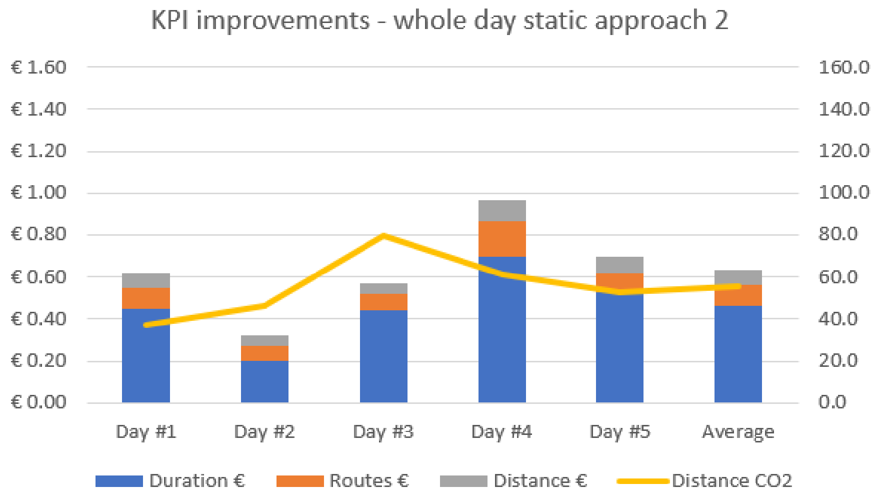


Figure 23: KPI improvements - whole day static approach 2

The outcomes of the static approaches are summarized in Table 46. The second approach performs better than the first approach in terms of both costs and sustainability.

	Costs	**Sustainability**
Static - approach 1	€0.53	46.2 $g\ CO_2$
Static - approach 2	€0.63	55.4 $g\ CO_2$

Table 37: Summary of the KPIs per order of the static approaches

7.2 Dynamic green labeling

In this subsection, the results of the dynamic green labeling approaches will be analyzed. *Duration*, *vehicles* and *distance* have the same definition here as for static simulation.

Approach 1 - whole day

In Table 38, the results are shown of the whole day optimization of the first approach within dynamic green labeling. Only whole day is shown instead of morning and afternoon plus evening since switching to another day part is possible. Separate day parts analysis will give an incorrect impression. In this approach, high-density areas receive a green label on the morning and afternoon four-hour time slots. Low-density areas get a green label on four-hour time windows that are offered in that area, see subsubsection 6.2.1 for a detailed explanation. Showing the results from separate day parts will be misleading. On average, the duration decreases, and therefore improves, by 8.9%, the amounts of routes decrease by 2.1% and the distance decreases by 7.4%. These results are compared to the results in the base case, where no adjustments are made. The results are also visualized in Figure 24.

	Day #1	**Day #2**	**Day #3**	**Day #4**	**Day #5**	**Average**
Duration	-11.7%	-8.9%	-9.2%	-6.5%	-8.1%	-8.9%
Routes	0.0%	-4.4%	-4.1%	0.0%	-2.1%	-2.1%
Distance	-9.3%	-7.3%	-7.1%	-5.1%	-8.2%	-7.4%

Table 38: Simulation results dynamic green labeling approach 1 whole day

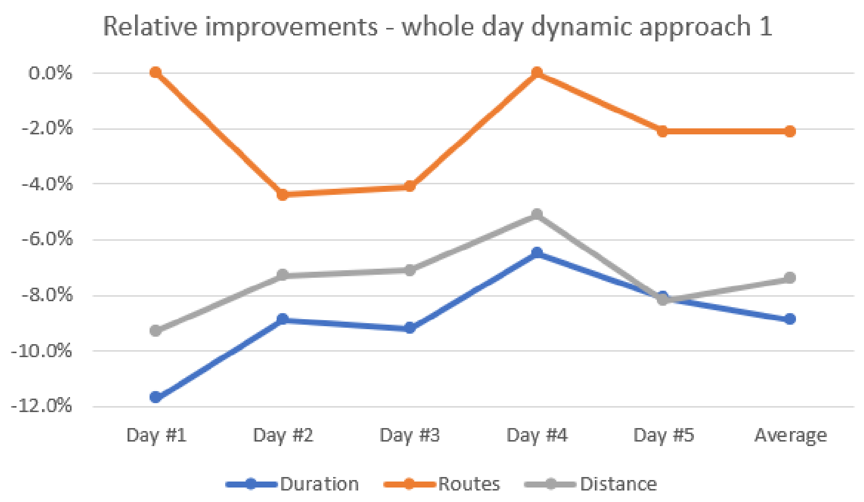


Figure 24: Relative improvements - whole day dynamic approach 1

Approach 2 - morning

In Table 39, the results are shown of the morning optimization of the second approach within dynamic green labeling. In this approach, high-density areas receive a green label on all four-hour time slots. Low-density areas get a green label on four-hour time windows that are offered in that area, see subsubsection 6.2.1 for a detailed explanation. On average, the duration decreases, and therefore improves, by 3.1%, the amounts of routes decrease by 2.7% and the distance decreases by 4.2%. These results are compared to the results in the base case, where no adjustments are made.

	Day #1	Day #2	Day #3	Day #4	Day #5	Average
Duration	-3.9%	-1.8%	-4.1%	-2.4%	-3.3%	-3.1%
Routes	-3.7%	0.0%	-4.8%	-3.2%	-2.0%	-2.7%
Distance	-3.6%	-4.4%	-5.1%	-4.4%	-3.6%	-4.2%

Table 39: Simulation results morning dynamic green labeling - scenario 2

Approach 2 - afternoon & evening

In Table 40, the results are shown of the afternoon plus evening optimization of the second approach within dynamic green labeling. On average, the duration improves by 3.5%, the amount of routes decreases by 3.2% and the distance decreases by 4.0%.

	Day #1	Day #2	Day #3	Day #4	Day #5	Average
Duration	-4.7%	-2.4%	-4.0%	-2.7%	-3.6%	-3.5%
Routes	-4.2%	0.0%	-4.6%	-3.9%	-3.3%	-3.2%
Distance	-3.6%	-3.4%	-4.9%	-4.7%	-3.4%	-4.0%

Table 40: Simulation results afternoon + evening dynamic green labeling - scenario 2

Approach 2 - whole day

In Table 41, the results are shown of the optimizations from the whole day of the second approach within dynamic green labeling. This means that the outcomes of the route optimizations are summed. On average, the duration improves by 3.3%, the number of routes decreases also by 3.0% and the distance decreases by 5.1%. The results are also visualized in Figure 25.

	Day #1	Day #2	Day #3	Day #4	Day #5	Average
Duration	-4.5%	-2.2%	-4.0%	-2.6%	-3.5%	-3.3%
Routes	-4.0%	0.0%	-4.8%	-3.6%	-2.9%	-3.0%
Distance	-3.6%	-3.7%	-4.9%	-4.6%	-3.4%	-5.1%

Table 41: Simulation results whole day dynamic green labeling - scenario 2

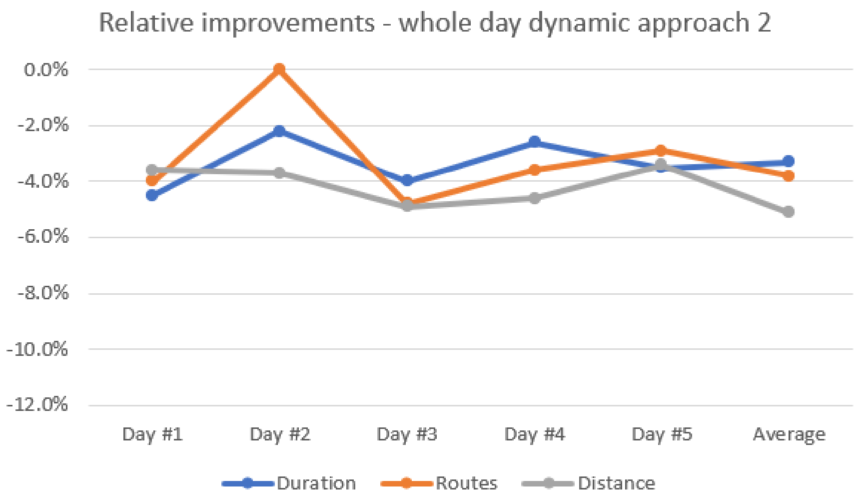


Figure 25: Relative improvements - whole day dynamic approach 2

7.2.1 KPIs

In order to compare both scenarios on costs, *duration*, used *vehicles* and *distance* are changed to monetary improvement, as described in subsection 4.2. With regard to sustainability, only *distance* is formulated in emissions. Only the results from the whole day are included since analyses by day part can give a wrong impression. In some approaches, orders will switch to another day part. In this way, comparisons are more reliable. Table 42 shows the improvement in costs per order. So on average, the costs per order become €1.03 lower in the first dynamic approach.

	Day #1	Day #2	Day #3	Day #4	Day #5	Average
Duration	€0.75	€0.32	€0.71	€1.04	€0.79	€0.72
Routes	€0.15	€0.13	€0.17	€0.28	€0.14	€0.17
Distance	€0.13	€0.14	€0.07	€0.24	€0.15	€0.14
Total	€1.03	€0.59	€0.95	€1.56	€1.08	**€1.03**

Table 42: Impact on costs from simulations - whole day dynamic approach 1

Table 43 shows the impact of distance on sustainability in terms of emissions. On average, 127.4 g CO_2 less is emitted per order.

	Day #1	Day #2	Day #3	Day #4	Day #5	Average
Distance	87.4 g CO_2	111.5 g CO_2	152.7 g CO_2	142.7 g CO_2	142.9 g CO_2	**127.4** g CO_2

Table 43: Impact on sustainability from simulations - whole day dynamic approach 1

The KPIs are also visualized in Figure 26, where the left axis indicated the monetary improvements per order and the right axis shows the improvement per order in terms of grams CO_2 per order.

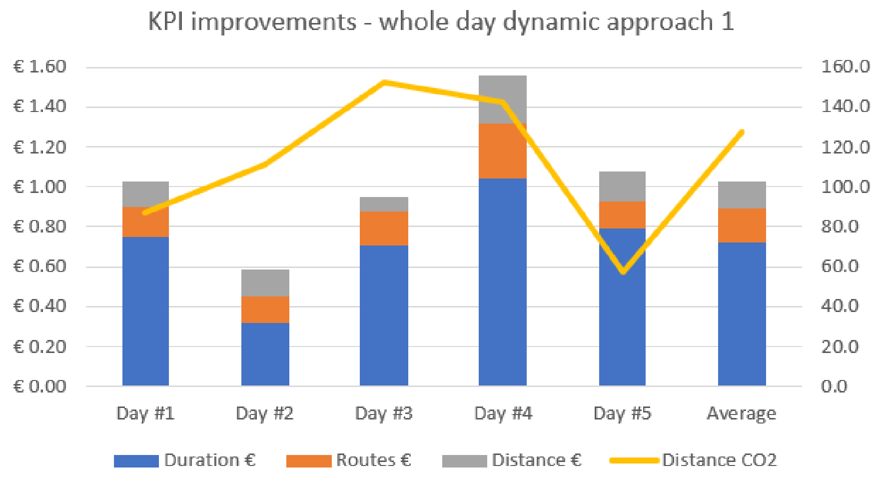


Figure 26: KPI improvements - whole day dynamic approach 1

Table 44 shows the improvement in costs per order. So on average, the costs per order become €0.61 lower in the second dynamic approach.

	Day #1	Day #2	Day #3	Day #4	Day #5	Average
Duration	€0.51	€0.21	€0.46	€0.55	€0.49	€0.44
Routes	€0.12	€0.07	€0.09	€0.12	€0.08	€0.10
Distance	€0.06	€0.06	€0.04	€0.10	€0.06	€0.07
Total	€0.69	€0.34	€0.59	€0.77	€0.63	**€0.61**

Table 44: Impact on costs from simulations - whole day dynamic approach 2

Table 45 shows the impact of distance on sustainability in terms of emissions. On average, 57.5 g CO_2 less is emitted per order. The KPIs for the second approach are also visualized in Figure 27.

	Day #1	Day #2	Day #3	Day #4	Day #5	Average
Distance	39.1 $g\ CO_2$	47.2 $g\ CO_2$	85.0 $g\ CO_2$	59.2 $g\ CO_2$	57.2 $g\ CO_2$	**57.5** $g\ CO_2$

Table 45: Impact on sustainability from simulations - whole day dynamic approach 2

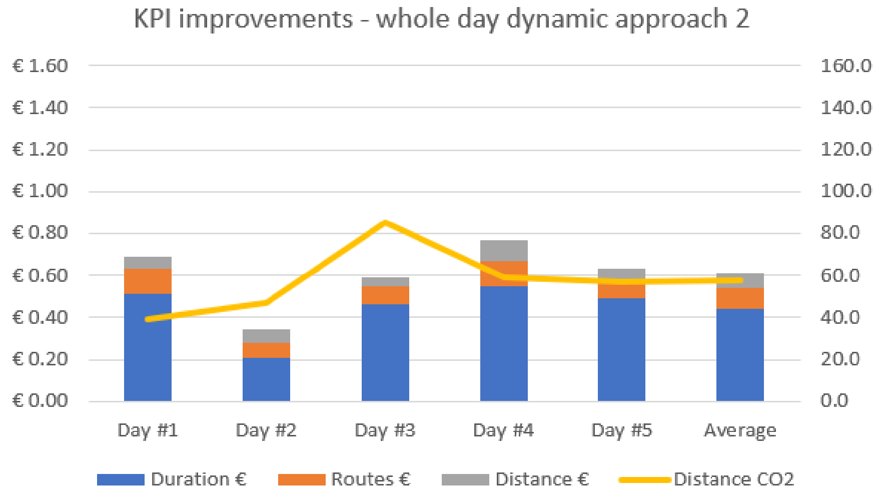


Figure 27: KPI improvements - whole day dynamic approach 2

The outcomes of the dynamic approaches are summarized in Table 46. The first approach performs better than the second approach in terms of both costs and sustainability.

	Costs	**Sustainability**
Dynamic - approach 1	€1.03	127.4 $g\ CO_2$
Dynamic - approach 2	€0.61	57.5 $g\ CO_2$

Table 46: Summary of the KPIs per order of the dynamic approaches

7.3 Other insights

In previous subsections, outcomes were compared within the different approaches, while it can also be valuable to compare the results per day. In terms of amount of order for the whole day, day 3 has the most orders, followed by day 1, day 4, day 5 and lastly, day 2. In Table 47, day 1 scores best in terms of relative improvements, followed by day 1, day 4, day 2 and lastly, day 5. This is an interesting outcome since this can imply that there is a relation between amount of orders and performance; the higher the total orders, the higher the effect of the green-labeling approach. Due to time constraints, further validation is not possible.

	Static 1	**Static 2**	**Dynamic 1**	**Dynamic 2**	**Average**	**Ranking**
Day 1	-3.3%	-4.2%	-9.3%	-3.6%	-5.1%	2
Day 2	-3.7%	-4.8%	-7.3%	-3.7%	-4.9%	4
Day 3	-4.6%	-5.5%	-7.1%	-4.9%	-5.5%	1
Day 4	-4.4%	-5.6%	-5.1%	-4.6%	-4.9%	3
Day 5	-3.4%	-3.8%	-8.2%	-3.4%	-4.7%	5

Table 47: Averaged optimization results per day

In Table 48, the percentage of orders that has its delivery location in a low-density area is shown. Day 3 has the highest share, followed by day 4, day 2, day 1 and day 5. Also here, a relation can be implied between the share of low-density delivery location and the effectiveness of green-labeling, although the relation is less clear than for amount of orders. Again, further validation is not possible due to time constraints.

	Low density	Ranking
Day 1	37.3%	4
Day 2	44.4%	3
Day 3	58.9%	1
Day 4	52.3%	2
Day 5	32.1%	5

Table 48: Relative orders in low-density areas

7.4 Summary of the results

Table 49 shows a summary of the KPIs for the different approaches simulated for the whole day. *Dynamic - approach 1* is the approach that leads to the highest improvement in terms of costs and sustainability per order. This is followed by *Dynamic - approach 2* that have better improvements in terms of sustainability compared to the number three, *Static - approach 2*. In terms of costs, number three performs slightly better than number two, but an improvement in sustainability weighs more. The lowest in the ranking is *Static - approach 1*.

	Costs	Sustainability	Ranking
Static - approach 1	€0.53	46.2 $g\ CO_2$	4
Static - approach 2	€0.63	55.4 $g\ CO_2$	3
Dynamic - approach 1	€1.03	127.4 $g\ CO_2$	1
Dynamic - approach 2	€0.61	57.5 $g\ CO_2$	2

Table 49: Summary of the KPIs per order of all the approaches

8 Discussion

In this chapter, the discussion will reflect on the main methods used in this research. With regard to the choice model and route optimization, it will be discussed how these methods can be applied in other fields and for other companies and also which extensions can be made for further research. Besides, the results from the previous section will be discussed.

8.1 Choice model

First of all, the beta estimates for the choice simulation resulting from the choice behavior model are based on historic data. The green labeling beta estimate results from another research. Using both data sources results in a good approximation of the reality but a downside is that it is not tested in real-life or via stated choice, which can be in the form of a survey. The MNL model relies on a few assumptions. First of all, customers never arrive at the same time, while in reality, this is not the case. This assumption is made in order to simplify the choice simulation.

The attributes of the choice models are the width of the time windows, the starting times and the costs since these are the only known attributes with regard to the characteristics of the time windows. A downside is that two attributes are related to each other: a wider time window has lower costs. What is not included are the characteristics of the customers themselves. When including for example attributes like age, location and gender, the choice model can be better estimated. This could have been estimated via a survey or with results from other research. This could also contribute to a better estimation of the smaller time windows, which are now difficult to address the right amount of orders to. However, this makes the model way more complex. Since all scenarios use the same values for the beta estimates, except for the difference with regard to the day part, the relative improvement is affected to a minimum.

In order to implement this choice model for a different situation, for example when a different data set from another company is used or the time windows will change within Crisp, the same attribute values can still be used. However, it is recommended to re-estimate the attribute values by changing the choice model in **??** to the specific situation. This ensures that the beta estimates are more reliable.

8.2 Route optimization

For this research, the route planning software Routigo is used, provided by Crisp. In this software, certain settings can be set differently, such as loading time, vehicle capacity and stop time. The base case and the different optimization approaches have the same settings in this research, which makes relative comparisons possible between the different approaches. It is assumed that another route planning software can also be used and should give the same relative results between the different approaches, but this is not tested.

An assumption that is made and has an important impact on the results is how is determined if a new order is within the same area as an already accepted order. This is based on the zip code of the orders, specifically on the first two digits of the zip codes. An advantage is that this way of distance determination includes indirectly the density of an area; zip codes with a high density have a smaller surface area in general. Assuming that these zip codes also have more orders, zip coding is an efficient way to determine proximity. Another advantage is that no distance matrix or such is needed. A downside is that real distance is not included.

With regard to the different choice simulation approaches, two approaches for the static scenario and two for the dynamic scenario are defined. The reason for this is is that first of all, static and dynamic are one of the main distinctions made in literature. Therefore, both are included in this research. Within static and dynamic green labeling, limited research is done and therefore every approach will contribute to scientific research. The four approaches are designed in such a way that together, they give a comprehensive view of the possibilities within static and dynamic. Since this research is bounded by time and other constraints, four approaches is set as the maximum.

8.3 Results

In this subsection, the results described in section 7 will be discussed. In general, the factors *duration*, *routes* and *distance* are related to each other. If a route can be planned more optimally, all three factors should go down. What is possible, and can be seen in the morning simulation of approach 1 of day 2, is that the amount of routes stays the same while the duration and distance decrease. When this happens, the routing software is able to optimize the routes since the orders are more clustered, but due to the volume capacity of the vans, fewer routes are not possible. The same happens in the morning simulation of approach 2 of day 2. The monetary and emissions values of the KPIs are based on values from Crisp and what is found online. More research on the exact values is recommended.

Since the results are first shown as relative improvements, it is possible that a lower relative improvement per order results in a higher KPI improvement per order for the same day and approach, compared to another day. This means that on that day, fewer orders or at least the total distance was lower. This is seen when comparing day 4 with day 1 for the *whole day static - approach 1*. Lower relative improvements lead to a higher KPI improvement for day 4.

Approach *dynamic - approach 1*, is the best approach mainly due to the nudging towards the morning and afternoon deliveries in high-density areas. In this way, the orders are better clustered in the less popular day parts and make more efficient routing possible. What should be tested is to which extent people are willing to switch to another day. A downside of the first dynamic approach is that it is more difficult to implement compared to a static approach. Since a static approach has easier decision rules, it can be worthy to first implement one of the two static approaches in order to test its effect in reality and use the feedback for the implementation of a dynamic approach.

Another point of discussion is that since the sustainability KPI is only linked to distance, it is possible that one approach performs better on one KPI and another approach better on the other KPI. This happens when comparing *static - approach 2* with *dynamic - approach 2*; the first leads to €0.63 less costs per order than the base case and 55.4 $g\ CO_2$ less per order while the latter leads to €0.61 less costs per order than the base case and 57.5 $g\ CO_2$ less per order. In this research, decreasing CO_2 emissions weights more than decreasing costs. Exact weights are left out of this research.

8.4 Managerial insights

As discussed in section 7, the routing performance can be improved most by implementing green labeling based on a dynamic approach where customers are nudged towards large time windows in less popular day parts. However, implementing this approach, and green labeling in general, comes with some challenges.

Capacity of morning and afternoon deliveries
In *dynamic - approach 1*, orders are nudged more towards the morning and afternoon orders. With regard to vehicle availability, this should not lead to a problem since the total amount of vehicles is determined based on the peak moments, which are the evening deliveries. What can lead to issues is the availability of the drivers. If drivers do not want to switch from evening shifts to morning or afternoon shifts, implementation could be difficult. Also, warehouse personnel should be flexible.

Implementation possibilities with regard to route planning software
For route planning software, it can be difficult to implement a time window offering based on dynamic labeling. It is important to discuss possibilities with the routing software company. With regard to static green labeling, this should be possible within every software package. The software company therefore also indirectly influences which labeling approach is suitable.

Effect on picking process
Shifting the deliveries more towards the morning and afternoon has besides the delivery capacity also impact on the picking process. While the picking process is left out of scope, it is worth mentioning here that by shifting deliveries to a certain part of the day, the picking process should also finish its process, in this case, earlier. Especially in the e-grocer sector, this has s substantial impact since fresh products are involved.

Communication of green labeling
When implementing green labeling, customers must be made aware of what green labeling exactly means and also what the effect is on the delivery process. The exact definition of green labeling should be communicated in the order process. For the dynamic approach, it can be distracting for customers when every time they place an order, a different time window has a green label. To what extent the customer is aware of this and what the impact is should be tested.

Implementation hybrid methodology to other companies and fields
The main research methods that are used in this research, a hybrid form of a choice model and route optimization, can also be applied to other companies and even other fields. Companies that also deal with time window offering can use the choice model in **??** where it has to be adjusted to its own parameters and according to its own data set. This has to be followed with the simulation steps described in section 6. A route software package or an own developed vehicle routing problem is needed. Adjustments with regard to the different approaches within static and dynamic are possible. While this study focused on steering customers towards green labels, other options are to steer customers towards other incentives. Before implementing this, this hybrid form of modeling can be used to estimate the possible effects of such a change before implementing.

The hybrid form of a choice model and route optimization can also be used in other fields. Every problem that has to deal with a form of choice modeling and some form of optimization, can use the outline presented in this research. Examples of optimization problems are scheduling problems and price optimization problems. These are common problems in the field of the airline industry, which is a good example of where this hybrid approach can be applied. Another interesting field where this hybrid form can be applied is public transport; when different attributes within choice behavior are estimated, the effect of changing ticket fares can be analyzed.

Sub question 5: What is the effect of green label time window offering on the routing performance?

Table 50 shows a summary of the KPIs for the different approaches simulated for the whole day. *Dynamic - approach 1* is the approach that leads to the highest improvement in terms of costs and sustainability per order. This is followed by *Dynamic - approach 2* that have better improvements in terms of sustainability compared to the number three, *Static - approach 2*. In terms of costs, number three performs slightly better than number two, but an improvement in sustainability weighs more. The lowest in the ranking is *Static - approach 1*.

	Costs	Sustainability	Ranking
Static - approach 1	€0.53	46.2 $g\ CO_2$	4
Static - approach 2	€0.63	55.4 $g\ CO_2$	3
Dynamic - approach 1	€1.03	127.4 $g\ CO_2$	1
Dynamic - approach 2	€0.61	57.5 $g\ CO_2$	2

Table 50: Summary of the KPIs per order of all the approaches

It can be concluded that dynamic green labeling has the most promising effect on routing performance, especially when customers are more nudged towards the largest time windows in the less popular day parts, which is the first approach. Other interesting insights are that green-labeling has likely a higher effect when the amount of orders is higher and when the orders are more in a low-density area. Further research on validation is needed.

The choice modeling and simulation and route optimization are both based on some assumptions. The results are not a full reflection of reality. The choice model is adapted to the time window offering of Crisp, which has the benefit that it reflects better a real process. A downside is that it is less generic. Also, only attributes with regard to the time window itself are included and not characteristics linked to the customer. With regard to route optimization, only one route software package is used and comparison between outcomes is therefore not possible. Also, only four approaches are tested, while there are countless options possible.

The research gap is reduced since this research contributed to increasing the knowledge on how to implement green labeling and what its effect is on relevant routing performance. Data from an e-grocer is used and a choice model was included. This combination was not used in scientific research before. More research on green labeling is needed, especially research on the implementation of green labeling in reality. However, green labeling is worth consideration for every e-grocer or even every parcel deliverer.

9 Conclusion and recommendations

First, the conclusion will emphasize the sub-questions and answers the main question in subsection 9.1. This is followed by recommendations on a practical and scientific level in subsection 9.2 and subsection 9.3, respectively.

9.1 Conclusion

This section will answer the main research question and therefore conclude this research. This research analyzed how to improve routing performance by implementing green labeling, taking into account customers' preferences, a new method within time window offering. In order to do so, route simulations were executed together with choice modeling simulations in order to take into account the preferences of customers, based on real data. Data from the e-grocer Crisp was used. What can be seen in practice is that customers become more aware of their ecological footprint and therefore are eager for more information about the sustainable impact of their choices or options to reduce that impact. Next to that, e-grocers try to lower their emissions by converting their fleet to a more electric one and improving their routing performance, in terms of sustainability and costs. Implementing green labeling is a promising tool to satisfy both the customer and the e-grocer.

In the literature study in section 2, three common choice models were studied, namely *fixed probability multinomial logit* and *general attraction*. The MNL model has the best fitting requirements and is therefore used in this research. Including a choice model in order to study the effect of green labeling based on real data is not done before and therefore contributes to the related academic theory.

The main research question is:

How to implement green labeling for an e-grocer's time slot offering in order to improve relevant routing performances, considering customers' preferences?

With the following sub-questions:

1. How can a last-mile delivery system for an e-grocer be defined?
2. What are the different options with regard to time window offering and demand management?
3. How can a choice model be estimated in order to investigate the impact of green labeling on slot choice?
4. How can a simulation study be set up in order to study the effect of green label time window offering on the routing performance?
5. What is the effect of green label time window offering on the routing performance?

Sub-question 1
A last-mile delivery system is the last part of the whole supply chain and consists of the following main components: pre-planning-, order- and planning system, warehouses, hubs, delivery vehicles and the customer. Last-mile delivery entails the transport from the hub or warehouse to the customer. The process starts with forecast and pre-planning, which provides capacity constraints with regard to e.g. deliverers and vehicles. The order process is important in terms of demand management since the options that are given to the customer have a great effect on the rest of the process. Also, customer satisfaction is influenced. When the customer placed an order, the planning process

can start, followed by the picking process and the order can be delivered to the customer from the warehouse or via a hub.

There are two KPIs when assessing routing performance, namely costs and sustainability, based on the factors duration, vehicles and distance, from which duration has the highest impact on costs and distance is the only factor impacting sustainability.

Sub-question 2
With regard to planning, it is important that the customer is at home during the delivery since a delivery failure is expensive due to the nature of the goods. Therefore, demand management within the e-grocer sector requires attended home delivery (AHD). The standard classification of demand management concepts is subdivided by pricing or time slot allocation incentives, which can occur in a static or dynamic environment. In a static environment, the possible time slot options are determined before the booking horizon starts based on characteristics such as the delivery location and time of day. A dynamic approach has the goal to serve time offering options individually for every customer request that occurs during the booking horizon. A relatively new form of incentive is green labeling, meaning that the labeled time slot will contribute to lower fuel consumption and emissions.

Besides a static or dynamic approach and used incentives, e-grocers can also differentiate their time window offering in terms of the number of time slots. Compared to other e-grocers, Crisp offers a wide variety of time slots and is, besides Picnic, the cheapest in terms of delivery costs. The time window offering in terms of window length and price occurs dynamically at AH and Jumbo. Crisp and Picnic deals with a static strategy, meaning that for a certain delivery location, the offering per day of the week is the same.

Albert Heijn is the only e-grocer in the Netherlands that makes use of green-labeled time slots. The green labeling occurs dynamically, but the greatest share of time slots with a green label are the time slots with a width of at least four hours. Green labeling is also seen at for example PostNL and e-grocers in other countries.

Sub-question 3
An MNL model will predict the probabilities of different options. It is multinomial since there are more than two possible discrete outcomes and the dependent variables are nominal. First, different beta estimates are calculated in Apollo with the data of Crisp. These beta estimates are used in the utility function without the green label attribute in order to estimate the choice probabilities of the MNL model. This is done for the morning deliveries, the afternoon and evening deliveries together and for all the orders on the same day. Thereafter, the beta estimate of applying a green label is included. With regard to the attributes for the whole day, the utility increases with 1.61123 when a time window becomes one hour bigger, 0.03134 when a time window starts one hour later and 0.04146 when the costs of the time window increase with one euro. The green-label attribute is binary, meaning that if a time window has a green label, the utility increases with 2.8600 and zero otherwise. This has resulted in larger time slots being chosen more often. In total, there is an increase of 9.4% in large time windows.

Sub-question 4
In static time slot management (STSM), the larger time windows receive a green label. Here, two approaches are defined. In the first approach, all the four-hour time windows receive a green label. In the second approach, the four-hour time windows are changed to five-hour time windows. Here, there is accounted for that in for example low density areas, only evening deliveries are performed.

Within dynamic incentive methods, there are two main directions; *intermediate vehicle routing problem (VRP) optimization* and *VRP optimization after final cut-off*. Since this research uses commercial routing software for the simulations, the green labeling should be determined on the locations of the already accepted orders and not on by means of intermediate routing optimization after each order. Therefore, the *VRP optimization after final cut-off* will be used. The green labeling will be done based on the zip codes of the accepted orders; customers receive adapted green labels when their location is within a certain proximity of already accepted orders for that day. There will be two dynamic labeling approaches simulated. The first approach has almost the same rules for high-density areas compared to the static situation. A difference is that this approach only places a green label on the morning and afternoon four-hour time window in order to nudge customers more towards the unpopular day parts For low-density areas, different rules will apply on which time window receives a green label. The second approach is similar, except that customers are not nudged more towards the morning or afternoon.

Sub-question 5
Dynamic - approach 1 is the approach that leads to the highest improvement in terms of costs and sustainability per order. This is followed by *Dynamic - approach 2* that have better improvements in terms of sustainability compared to the number three, *Static - approach 2.* In terms of costs, number three performs slightly better than number two, but an improvement in sustainability weighs more. The lowest in the ranking is *Static - approach 1.*

The choice modeling and simulation and route optimization are both based on some assumptions. The results are not a full reflection of reality. The choice model is adapted to the time window offering of Crisp, which has the benefit that it reflects better a real process. A downside is that it is less generic. Also, only four approaches are tested, while there are countless options possible.

It can be concluded that dynamic green labeling has the most promising effect on routing performance, especially when customers are more nudged towards the largest time windows in the less popular day parts, which is the first approach. An important note is that these results hold for this specific case study. It is assumed that the results are generalizable, but to what extent is not tested and recommended for further research.

Main research question
Now all sub-questions are answered, the main research question can be answered:

How to implement green labeling for an e-grocer's time slot offering in order to improve relevant routing performances, considering customers' preferences?

First, choice behavior has to be modeled in order to gain information on what the potential is of customers switching to a green-labeled time slot. This can be done via a choice model based on data, as in this research, or via a stated choice survey. Another option is to test this in real-life via a new version of the delivery app. In this research, there is an increase of 9.4% in large time windows according to the choice model.

Results show that all approaches lead to an improvement in costs and sustainability. The most optimal approach, *Dynamic - approach 1*, leads to an improvement of €1.03 and 127.4 $g\ CO_2$ per order. This is based on an average improvement of 8.9% of the duration, 2.1% of the routes and 7.4% of the distance. This does not mean that other incentives, for example pricing, are of less use.

Pricing can still be an interesting tool from a marketing perspective; it can be used as a tool for market segmentation and for increasing profit. It can be concluded that steering customers towards green labels is a promising, environmentally friendlier option for e-grocers and an important topic for further research.

9.2 Recommendations for Crisp

The results of this thesis lead to several recommendations specifically for Crisp, of which some are also applicable to other e-grocers. First of all, the choice simulation is done manually in Excel in order to determine which time window would get a green label, following one of the dynamic approaches. If Crisp wants to implement dynamic green labeling, a consult with Routigo or other route planning software companies is needed in order to discuss what the options are. What is also recommended, is to test how many customers will really switch to a green time window. This can be done in the app test environment of Crisp consisting of employees of Crisp as users. Especially static approaches can easily be implemented and can serve as a real-life choice behavior experiment. Therefore, it is recommended to first implement one of the two static approaches, followed by one of the dynamic approaches. Also, it is important to determine which of the approaches fits the wishes and constraints bests. If it is not possible to serve more customers in the morning and afternoon, a good second option can be to only implement the second static approach; since this approach performs only a bit lower than the second dynamic approach, it can be considered to opt for static due to its ease of implementation. When implementing a dynamic approach, it is recommended to implement green labeling based on zip codes since this is easy to implement and leads to lower computation time.

Another recommendation resulting from observations from how competitors have set up their time windows is to increase delivery costs and decrease the number of time windows and especially one-hour time slots. Besides, it is advised to explore options with regard to subscription models for the delivery costs.

9.3 Recommendations for future research

This research used zip codes in order to determine if deliveries are in the same area and if the area is within a low or high-density area. For the latter, zip coding is the most useful method, while for the distance between deliveries, a lot of other options are possible, for example, numerous clustering methods can be used or grouping based on Euclidean distance. Even though the methods used in this research are fitting the requirements, using other grouping methods could be interesting to test and is therefore recommended.

For the choice model, order data for almost a whole year was used, which gives a good representation of the real attribute values. The beta values of the attributes start time, width and costs of time window were estimated. A downside of this approach is that the attributes are related to each other, which makes it more difficult to estimate. This is visible in the validation of small time windows, where the results from the choice model are very different than the real values. Another option could be to include more customer-specific attributes, such as age, gender and zip code. A downside is that the choice model becomes more complex.

With regard to the days used for the route optimization, five different weekdays are used. It is interesting to use more days and also include weekend days. This could contribute to the reliability of the results. This study only focused on four different approaches, either static or dynamic, while countless approaches are possible. One interesting approach to test would be a static approach where green labels are put only on the morning and afternoon time windows. In this way, the nudging towards these day parts, what happens in *dynamic - approach 1*, can be approximated in a static way. Furthermore, only different approaches are tested, while it could also interesting to see what the effect is of a different definition of a low- or high-density area and what the effect is when the attributes of the choice model have a different value. With regard to route optimization, the raw data set of Crisp is used, which makes generalization of results less possible. It is recommended to do the route optimization experiments by means of an order generator, based on attributes of a data set such as amount of products, delivery area and slot choice.

Bibliography

Agatz, N., Campbell, A., Fleischmann, M., and Savelsbergh, M. (2011). Time slot management in attended home delivery. Transportation Science, 45(3):435–449.

Agatz, N., Campbell, A. M., Fleischmann, M., Van Nunen, J., and Savelsbergh, M. (2013). Revenue management opportunities for internet retailers. Journal of Revenue and Pricing Management, 12(2):128–138.

Agatz, N., Fan, Y., and Stam, D. (2021). The impact of green labels on time slot choice and operational sustainability. Production and Operations Management, 30(7):2285–2303.

Campbell, A. M. and Savelsbergh, M. (2006). Incentive schemes for attended home delivery services. Transportation science, 40(3):327–341.

Campbell, A. M. and Savelsbergh, M. W. (2005). Decision support for consumer direct grocery initiatives. Transportation Science, 39(3):313–327.

Chevalier, S. (2021). Online supermarket revenue in the netherlands from october 2015 to june 2021. `https://www.statista.com/statistics/659321/online-supermarket-revenue-in-the-netherlands/`.

Hays, T., Keskinocak, P., and López, V. M. d. (2005). Strategies and challenges of internet grocery retailing logistics. In Applications of supply chain management and e-commerce research, pages 217–252. Springer.

Insider, B. (2021). De online boodschappenmarkt groeit hard en daar profiteren vooral de kleine supers van – picnic stoot jumbo van de 2e plaats. `https://www.businessinsider.nl/online-boodschappenmarkt-ah-jumbo-picnic/`.

Klein, R., Neugebauer, M., Ratkovitch, D., and Steinhardt, C. (2019). Differentiated time slot pricing under routing considerations in attended home delivery. Transportation Science, 53(1):236–255.

Köhler, C., Ehmke, J. F., and Campbell, A. M. (2020). Flexible time window management for attended home deliveries. Omega, 91:102023.

Macioszek, E. (2017). First and last mile delivery–problems and issues. In Scientific and technical conference transport systems theory and practice, pages 147–154. Springer.

Mackert, J. (2019). Choice-based dynamic time slot management in attended home delivery. Computers & Industrial Engineering, 129:333–345.

Mangiaracina, R., Perego, A., Seghezzi, A., and Tumino, A. (2019). Innovative solutions to increase last-mile delivery efficiency in b2c e-commerce: a literature review. International Journal of Physical Distribution & Logistics Management.

Nomura, N. and Akai, M. (2004). Willingness to pay for green electricity in japan as estimated through contingent valuation method. Applied Energy, 78(4):453–463.

Ortec (2022). Klanten: Albert heijn. `https://ortec.com/nl-nl/klanten/albert-heijn`.

Rana, J. and Paul, J. (2017). Consumer behavior and purchase intention for organic food: A review and research agenda. Journal of Retailing and Consumer Services, 38:157–165.

Strauss, A., Gülpınar, N., and Zheng, Y. (2021). Dynamic pricing of flexible time slots for attended home delivery. European Journal of Operational Research, 294(3):1022–1041.

Treacy, M. and Wiersema, F. (1993). Customer intimacy and other value disciplines. Harvard business review, 71(1):84–93.

Truong, H. (2022). Understanding the supply chain process. `https://altametrics.com/supply-chain-management/supply-chain-process.html`.

van den Hooven, B. (2018). 'groen bezorgslot werkt bijna net zo goed als goedkoop bezorgslot'. `https://twinklemagazine.nl/2018/05/niels-agatz-last-mile-delivery/index.xml`.

van Loon, D. (2022). Online supermarktbestedingen normaliseren, maar picnic groeit door. `https://www.foodpersonality.nl/online-supermarktbestedingen-normaliseren-picnic-groeit/`.

van Mil, J. (2022). Improving last-mile transport planning by nudging customer delivery windows based on a dynamic spatial customer grouping: Simulating nudging of customer delivery windows at order creation.

Williams, A. (2022). Commission's 2021 proposal for new van co2 targets. `https://www.transportenvironment.org/challenges/road-freight/vans/`.

Yang, X. and Strauss, A. K. (2017). An approximate dynamic programming approach to attended home delivery management. European Journal of Operational Research, 263(3):935–945.

Yang, X., Strauss, A. K., Currie, C. S., and Eglese, R. (2016). Choice-based demand management and vehicle routing in e-fulfillment. Transportation science, 50(2):473–488.

Zhang, D. J., Dai, H., Dong, L., Qi, F., Zhang, N., Liu, X., Liu, Z., and Yang, J. (2020). The long-term and spillover effects of price promotions on retailing platforms: Evidence from a large randomized experiment on alibaba. Management Science, 66(6):2589–2609.