Evolution of Consumer Preferences in Last-mile Delivery Methods and the Impacts on City Logistic Freight Traffic

a simulation study

J.E. van Vliet

Master Thesis

Keywords:
Last-mile, consumer preference, freight logistics, e-commerce, system dynamics

TU Delft

significance
quantitative research

Evolution of Consumer Preferences in Last-mile Delivery Methods and the Impacts on City Logistic Freight Traffic

a simulation study

by

J.E. van Vliet

to obtain the degree of
Master of Science Transport, Infrastructure and Logistics
at the Delft University of Technology and Significance,
to be defended publicly on August 31, 2023 at 11:00 AM.

Student number:	4600290	
Project duration:	February 13, 2023 - August 31, 2023	
Thesis committee:	Dr. ir. G. Homem de Almeida Correia,	TU Delft, chair
	Dr. B. Atasoy,	TU Delft, supervisor
	Dr. ir. M. A. de Bok,	Significance

Preface

This master's thesis marks the end of my student time at TU Delft. With a general interest in technology, I discovered after a few years that mobility and transport are the topics that I am really enthusiastic about. Therefore, this thesis is the final work to obtain a master's degree in Transport, Infrastructure and Logistics.

First of all, I would like to thank my graduation committee, as they were genuinely a pleasure to work with. As this project started with only a general idea, Gonçalo Correia and Bilge Atasoy helped me in defining the research topic and finding a direction close to my interests. Their feedback and ideas during discussions or in comments provide highly appreciated guidance. Additionally, I would like to add that the speed at which you found time to help me was remarkable.

I also want to especially thank Michiel de Bok, who supervised me on a weekly basis. With those regular meetings, it was very pleasant to discuss large and small steps, which helped to easily make progress each week. Furthermore, I felt very welcome at Significance, where everyone was very kind to help me or generate a good atmosphere by discussing the latest Tour de France or Scorito standings.

Lastly, it would be great if the research helps in the further development of both last-mile models and establishing more sustainable solutions to fulfil our untempered demand.

J.E. van Vliet
Delft, August 2023

Summary

The market for online shopping (e-commerce) has been growing over the years, becoming more common thanks to the spread of laptops, tablets and smartphones (Morganti et al., 2014). All age groups adopt online shopping, and even elders are engaging in this trend (CBS, 2022). The fulfilment of these orders comes with logistical challenges, especially last-mile delivery is a crucial success factor and an influential aspect in the consumer purchase decision (Nguyen et al., 2019). Last-mile services generate problems, mainly in inner cities, because the rising number of delivery vehicles causes traffic congestion and air pollution (Deloison et al., 2020). New delivery methods must be implemented to deal with the growing parcel demand and to make the last-mile more sustainable. However, estimating the impact of a different logistics system is complex, as it is highly subjected to the adoption of new delivery options (Maltese et al., 2021). Hence, it is essential to understand consumer preferences for last-mile delivery.

The objective of this research is to develop a method that simulates to what extent consumer preferences evolve over time due to the performance of conventional and emerging delivery methods in a spatial model. For that purpose, influential factors for consumer preference development are discussed and modelled, with the added novelty of gathering empirical data on the direction and magnitude of those factors at multiple time points. Furthermore, the evolution of the supply of delivery methods is replicated. This approach allows the exploration of the complex last-mile system where multiple delivery methods compete and complement each other.

The parcel last-mile to consumers is currently dominantly performed with human-driven vans (Boysen et al., 2021). Although van delivery has large disadvantages such as low efficiency, reliability on the consumer being at home and the reliance on and cause of traffic, other delivery methods are not yet disrupting the last-mile environment. Electrification is used to reduce the environmental impact of vans (Buldeo Rai et al., 2021), and cargo bikes replace vans in more and more urban areas. In Europe, pick-up points are a well-known alternative for at-home delivery (ACM, 2020), and the addition of automated lockers boosts the flexibility for consumers (Boysen et al., 2021). However, still, only 18% of the parcels in the Netherlands are delivered to a self-collection point (ACM, 2020), and they are mostly used for returning parcels (Weltevreden, 2008).

New delivery methods are under development, and real-world test cases are being executed with drones and droids (Boysen et al., 2021). Both these delivery methods use automation to reduce the operational costs of the carrier. Thereby, the prediction is that drones and droids can offer fast and flexible delivery that is convenient for the consumer (Leon et al., 2023). Furthermore, drones open up new possibilities as they are not dependent on road infrastructure. However, both delivery methods still need to overcome technical and especially regulatory barriers. Another disadvantage is the low capacity of these modes. Consequently, in many cases, consolidation centres will be needed (Mohammad et al., 2023). Crowdshipping is another development that can reduce the transport burden of carriers, as it uses the crowd to move or deliver parcels during their own travel (Gatta et al., 2018) (Punel et al., 2018). A crowdshipper gets remuneration for their work/effort (Buldeo Rai et al., 2021), which provides the carrier with a highly flexible and scalable workforce. In this case, drawbacks could be inducing traffic, safety concerns and the dependency on the willingness to act as a crowdshipper (Tapia et al., 2023). New innovations are also introduced for unattended home delivery, which eliminates the problem of consumers not being at home (Boysen et al., 2021).

In this thesis, it is chosen to implement van, self-collection, crowdshipping and drone delivery. The main reason for this choice is that all four delivery methods have distinctive characteristics that give the consumer a unique experience during the handover.

Consumers evolve their preferences for delivery methods based on multiple factors, and this research applies the theory of perceived characteristics in combination with a logit formula to derive the probabilities of consumers preferring a specific delivery method (Hepp, 2018) (Shabbir et al., 2017) (Walker & Ben-Akiva, 2002). To obtain an indication of the perceived characteristics, four attributes are used that score the operational performance of each delivery method. Delivery costs and speed are the most used in comparable research. Additionally, a reliability attribute is incorporated that can also describe failed deliveries. Furthermore, as only self-collection delivery requires consumers to pick up their parcel, an attribute is added for the distance consumers need to travel.

Secondly, the theory of word of mouth (WoM) is implemented, as it describes the process of consumers evaluating a service and communicating that experience with other consumers (De La Torre et al., 2019). Herefore, the Bass diffusion model is used, which formulates the growing group of people that can spread WoM due to their experience and a shrinking group of consumers that can still adopt the service, which determines the corresponding amount of new adopters at a certain point in time because of WoM. Familiarity is another factor that describes the dynamics of consumer adoption of services. According to this theory, consumers build trust and loyalty towards earlier used services or products, making them less likely to opt for an unknown delivery method.

To be able to simulate the evolution of consumer preferences and the operational fulfilment of the deliveries, a hybrid model is developed that combines System Dynamics (SD) modelling and Agent-Based Modelling (ABM). The SD will be used to model the preference evolution of consumers over time. The ABM simulates the carriers delivering the parcel from the depot to the receiver and the evolution of performance characteristics of each delivery method. In Figure 1, it can be seen how the hybrid model is connected.

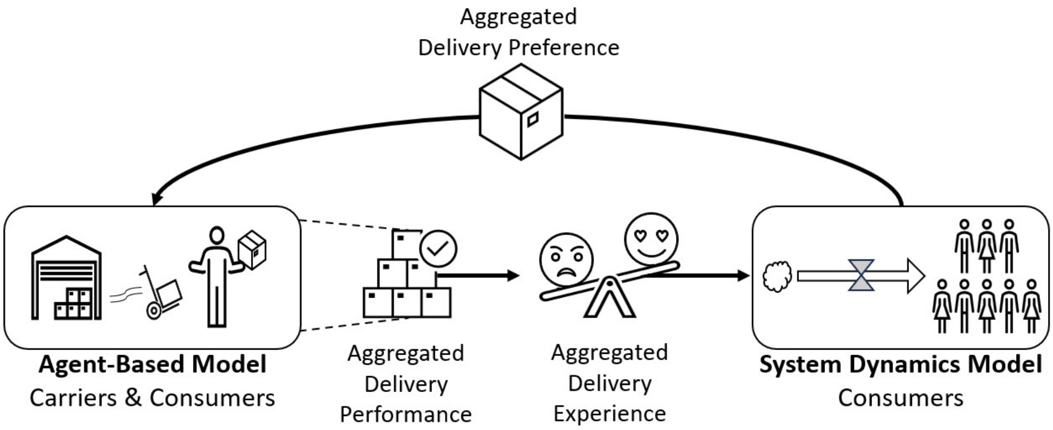


Figure 1: Schematic overview of the connection between the SD and ABM model

There is a feedback loop between the ABM and SD models; thus, both models will react to each other's output. At the simulation start, an estimated aggregated preference probability for each delivery method is set. That will result in a number of requests for each delivery method for all zones. Those requests are the input for the ABM model, where all carriers will fulfil the requests. That will constitute the operational performances of all parcel delivery requests, which are aggregated to an average performance per delivery method. Subsequently, the SD model will estimate the new aggregated delivery preferences based on the delivery method performance, WoM and familiarity. Additionally, via SD, carriers can adapt the operation of each delivery method, for example, the number of drones, based on the demand for that delivery method. With this structure, consecutive interactions can be simulated, and by that, data on consumer preferences and the delivery method operation can be obtained for each iteration and at a final horizon.

A case study is performed on the Dutch province of South Holland, a highly urbanised area with 3.6 million inhabitants, various large cities and a well-developed transport network. The case study simulates five years, in which an iteration resembles a quarter of a year. Each year the parcel demand

grows with a constant growth factor. For this case study, three different scenarios are explored: 1) Current state: van and self-collection delivery; 2) Addition of crowdshipping; 3) Innovation: also including drone delivery.

Before the scenario analysis, the model is improved by performing a calibration of the initial self-collection distribution and sensitivity analysis. It is established that the model is most sensitive to the beta weights of the costs and reliability attributes and that the impact of WoM and familiarity are moderate. The developed algorithm for the evolution of the self-collection points is very sensitive to the threshold values, which strongly affect the performance of self-collection delivery. The developed simulation model is verified and validated with multiple tests, and it is concluded that the model results can be used for the purpose of this study.

The simulation results estimate that in all three scenarios, van delivery will be the delivery option that most consumers prefer; see Table 1. The service level that van delivery can offer is constantly high across all zones. The growth in parcel demand leads to an enormous increase in the number of trips that have to be performed per day. The addition of other delivery methods can significantly reduce the demand for van delivery. Although the performance scores of the other delivery methods are on average higher, thus performing worse, many consumers will still choose those delivery methods.

Table 1: KPIs scenarios

Indicator	**Method**	**Scenario at t=20** [quarter]		
		1	**2**	**3**
Market share	Van	69%	53%	39%
	Self-Collection	31%	25%	18%
	Crowdshipping		22%	15%
	Drone			28%
Trips per day	Van	153277	116497	84181
	Self-Collection	8052	6890	5053
	Crowdshipping		23169	17729
	Drone			36324
Vehicle kilometers per day	Van	108793	96017	83205
	Crowdshipping		45590	36785
	Drone			195858
CO2 emissions per day	Van	22904	19916	16940
	Crowdshipping		3047	2462
	Drone			0
Number of Self-Collection points		2250	1956	1508
Total capacity Self-Collection points		44500	39000	30000
Number of Drones				994

Self-collection delivery shows strong spatial dependency and can be a competitive delivery method in dense urban areas. With the addition of crowdshipping and drone delivery, the results show that fewer consumers will prefer self-collection. Because of those lower demands, the distribution of self-collection becomes scarcer, and the total capacity after five years, $t = 20$, is around 13% and 31% lower for scenarios 2 and 3, respectively. Indicating a negative reinforcing feedback between the demand and supply.

Crowdshipping offers, like van delivery, a constant delivery performance. However, the service is generally of mediocre quality, thus making this delivery method the least preferred. The introduction of drone delivery has a large impact on the predicted market shares of the other delivery methods. Drones do perform well around the depots, and therefore a lot of drones are needed to fulfil the demand. However, the performance in the outskirts of the study area is generally low.

Finally, the large yearly growth in parcel demand results in higher vehicle kilometres and CO2 emissions, regardless of the scenario. As can be seen in Table 5.14, the total demand almost doubled in five years. The introduction of crowdshipping reduces the demand for van delivery and self-collection, yet it will result in more vehicle kilometres that need to be made on the same road infrastructure as van delivery. However, crowdshipping produces less CO2 per km, and therefore the CO2 emissions are very comparable with scenario 1. The high CO2 emissions in scenario 2 can be explained by vans becoming more efficient when the demand is higher. With increasing parcel demand, the number of zones that need to be visited increases slightly, and those transports contribute to the largest share of vehicle kilometres and emissions. Hence, the positive effect of the substitution of van and self-collection delivery by crowdshipping is partly reversed. This can also be seen in Table 5.15, where the average kg CO2 emission per parcel is shown. In all scenarios, this number improves. The main reason herefore is vans becoming more efficient due to the absolute increase in parcel demand delivery by van. Scenario 3 results in the lowest CO2 emissions due to drone delivery being emission-free and results in 15% less CO2 emission than scenarios 1 and 2. Additionally, the van vehicle kilometres and emissions within zones are much lower than in the other scenarios, which reduces the last-mile burden in urbanised areas and on low-capacity roads.

Table 2: Average C02 emissions per parcel [kg/parcel]

Indicator	Scenario 1	Scenario 2	Scenario 3
t = 0	0.160	0.159	0.159
t = 20	0.103	0.103	0.087

It can be concluded from the results that the preference probability of consumers strongly responds to the perceived performance and WoM effect. As the offered services from each delivery method are comparable to a great extent, it is expected that each delivery method will generate a significant market share. However, to effectively reduce the vehicle kilometres and the emissions & costs that come with last-mile parcel delivery, governments and carriers should explore ways to stimulate consumers in choosing a sustainable delivery option. This study indicates that large investments are needed to provide the infrastructure (number of self-location points) or vehicles (drones) to handle the demand. To boost a shift from van to other delivery methods, even higher investments are needed.

The new model has three main contributions. The interaction between the demand and supply of several parcel delivery methods in a spatial model allows the exploration of the complex last-mile system where multiple delivery methods compete and complement each other. The second contribution is the recognition of stochasticity in the delivery operation. This supports future research into the interaction between consumers and the reliability of delivery methods. Thirdly, this modelling approach provides insight into the direction and magnitude of various factors that take place, with the added novelty of gathering that empirical data at multiple time points.

Various improvements could advance the developed simulation model. By changing the model from an aggregation level over the entire study to a lower level, like a municipality, a zone or even an agent, it will be possible to explore spatial differences closer. An additional advantage is that a lower-level SD implementation offers the possibility of differentiating the decision rules. For example, the beta weights for each attribute could be adapted based on the socio-demographic characteristics of a consumer or zone. Furthermore, a moving model memory can account for consumers' recollections of multiple deliveries. Fourthly, the WoM effect is based on the Bass diffusion model, which is intended for innovative products or services; thus, not exactly applicable to the scenarios of this study. Lastly, the models for the operation of the delivery methods can be improved in various ways. For example, the drone delivery model is quite simple and real operations should take much more constraints into account, or a dynamic willingness to crowdshipping can be modelled.

Future research could improve the understanding of consumer preferences for last-mile delivery solutions. First of all, available data sets on consumer preferences for different delivery methods are scarce, and research has, for example, not yet settled on best practices. Higher predictive accuracy could be achieved with this and other models when new data collection efforts provide insight into consumer preferences for delivery attributes with large sociodemographic groups. Thereby, many stated

preference studies neglect the possibility of a failed delivery, where this study shows to be noticeably sensitive to the reliability attribute as it has a strong dynamic relation between consumer preferences and operational performance. Therefore, it is suggested to include reliability in further research. Additionally, future research could explore how consumers take environmental considerations into account when deciding on a delivery method.

List of Abbreviations

AHD	Attended Home Delivery
AB	Agent-Based
ABM	Agent-Based Modelling
ACM	Autoriteit Consumer & Markt
APL	Automated Parcel Locker
B2C	Business to Consumer
C2C	Consumer to Consumer
CC	Consolidation Centre
CLD	Causal Loop Diagram
DOI	Diffusion of Innovation
KPI	Key Performance Indicator
MASS-GT	Multi-Agent Simulation System for Goods Transport
RP	Revealed Preference
R&D	Research and Development
RUM	Random Utility Model
SD	System Dynamics
SP	Stated Preference
SFD	Stock and Flowd Diagram
UAV	Unmanned Aerial Vehicle
UGV	Unmanned Ground Vehicle
UHD	Unattended Home Delivery
VKMs	Vehilce kilometers
WoM	Word of Mouth
WTP	Willingness to Pay
WTW	Willingness to Work

Contents

List of Figures

List of Tables

1

Introduction

In this thesis, a simulation model is developed to capture and explore the interconnections between developing delivery methods and consumer preferences for delivery methods. Consumers' reaction to the performance and availability of the delivery methods is simulated, next to knowledge progress via word of mouth and familiarisation. The developed model uses a system dynamics model to simulate the evolution of consumer preferences and the last-mile delivery methods at an aggregate level, while an agent-based model provides input on the operational performance of all delivery methods. In this chapter, the research problem, scope and objectives are defined.

1.1. Context

The Paris Agreement aims to reduce CO2 emissions by at least 50% in 2030 (Punte et al., 2019). However, freight emissions are still increasing, and with an estimated tripling of global transport demand by 2050, a complex problem arises. A doubling of carbon emissions is expected if we continue in our current manner. Last-mile delivery is part of the supply chain where the consequences, like pollution, congestion and noise, for the freight logistics have the largest impact (ITF, 2022). Five problems that must be dealt with can be identified (Boysen et al., 2021): 1) increasing volume; 2) sustainability; 3) costs; 4) time pressure; 5) an ageing workforce. In addition, DHL marks the last-mile as the "least predictable part of the entire journey" ("Self-Driving Vehicles in Logistics. A DHL perspective on implications and use cases for the logistics industry", 2014). The logistics sector seeks to innovate the last-mile problem with emerging delivery methods, new delivery locations and automation. Generally used delivery vans are being electrified or could potentially be replaced with other modes, like drones, bike couriers and droids (Joerss et al., 2016). Also, crowdshipping and locker points can provide effective solutions. These innovations could provide hard benefits, like reduction of costs, and soft benefits, like competitive differentiation and service enhancements between parcel carriers (Hepp, 2018).

However, the consumers' acceptance of new and unfamiliar delivery options is crucial for the success of these innovations (Buldeo Rai et al., 2019) and could thus provide a barrier to innovation adoption. Consumers judge these services on their attributes (Buldeo Rai et al., 2021), yet consumers also rely on word of mouth (Wong & Sheng, 2012) and familiarity (Fandos Herrera & Flavián Blanco, 2011) to develop their preferences. As the estimation of the impact of a different logistics system is highly subjected to the adoption of new delivery options (Maltese et al., 2021), it is complex to predict which last-mile innovations provide the most sustainable solution.

1.2. Problem Definition

Consumer preferences play a major role in forecasting the demand for parcel deliveries, as these preferences affect the consumer choice of the parcel delivery method. Consumers order increasingly more packages and are highly sensitive to delivery characteristics (Joerss et al., 2016). At the same time, to reduce the carbon footprint and the impact on city logistics, new services and technologies are adapted to make last-mile deliveries more efficient (Deloison et al., 2020). Predicting how these consumer preferences will evolve is complicated due to the dynamics introduced by an uptake in demand & capacity and price setting for new alternatives. Furthermore, knowledge about the mix of delivery methods is needed, as each last-mile concept has its own strengths and weaknesses (Boysen et al., 2021).

Most studies analyse consumer preferences for last-mile delivery, including preferences for new services, using conventional Revealed Preference (RP) or Stated Preference (SP) data, see chapter 3. From this, we know that delivery price and time are dominant factors in the preferred delivery. And although reliability and ease of use are not often considered in forecasting models, they certainly influence consumer preferences over time (Nguyen et al., 2019). However, the reliability and ease of use of the new last-mile delivery options heavily depend on demand levels. In many cases, these systems have limited capacity (e.g. number of parcel lockers, number of cargo bikes/delivery robots available). The introduction of these services goes hand-in-hand with user experiences that will affect consumers' preferences for their preferred delivery option.

Yet research, especially empirical-based, on consumer preferences is scarce. Likewise, the integration of multiple different delivery methods in one consumer preference model is limited. The problem with the conventional RP or SP survey data is that they cannot measure future preferences (RP-data) or consider the dynamics between demand levels and capacity and the price setting of new solutions (SP-data). This creates a gap in how to determine the evolution of user preferences while taking into consideration the dynamics in the delivery system (e.g. price, capacity, reliability). Closing that gap could help in creating long-term viable delivery solutions from a logistics and commercial side.

1.3. Scope

This study will develop a method to simulate to what extent consumer preferences evolve over time due to the performance of conventional and emerging delivery methods. Hereby, the delivery preferences and the subsequent delivery choice depend on operational performance and the consumer's attitude towards delivery methods. Both inputs are dynamic over time, and therefore it is assumed that consumer delivery preferences evolve. By developing such a model, this study contributes in two ways: 1) it will create more understanding of consumer preferences in a setting with multiple delivery methods; 2) it helps to understand long-term viable delivery solutions.

A will be developed model that quantifies the evolution of consumer preferences for last-mile deliveries, thus, Business to Consumer (B2C) delivery. Hereby, the focus will be on four distinctive delivery methods: 1) van delivery; 2) delivery to self-collection points; 3) crowdshipping delivery; 4) drone delivery. The operation of new delivery methods or options will be based on literature and will not be self-developed. The preferred delivery from the company perspective is not considered for the evolution modelling. To estimate parcel demand, an existing freight logistics model, MASS-GT, is used; thus, this thesis will not focus on the parcel demand. Furthermore, the current vehicle, route assignment and emissions model will be used without improving assignment strategies. As MASS-GT is developed for the province of South Holland in the Netherlands, this research will be placed in a Dutch context. This study will explore the interactions between different factors in the last-mile environment. Thereby, the main focus is creating an understanding of how different dynamics affect each other. Thus, the simulation model is not developed for exact forecasting.

1.4. Objectives & Research Questions

The objective of this research is to describe the evolution of consumer preferences in last-mile delivery methods in a future state and the impact of those preferences on logistic freight traffic. The evolution will be simulated by consumers experiencing the operation of delivery methods, becoming aware of new delivery methods via, for example, word of mouth and the growth or decline in the operating capacity of delivery methods. This thesis aims to fill the knowledge gap about the dynamic interaction between consumer preferences and last-mile operation. It helps to understand how consumers value different parcel delivery methods and how consumers adapt that valuation over time. These evolving preferences affect the operation of last-mile logistics and can thus provide insight into the viability of a delivery method in an integrated system. Results will be measured in market shares of the different last-mile methods and by vehicle kilometres and emissions.

To achieve this objective, the following main research question should be answered:

What is the impact of evolving consumer preferences for last-mile delivery on parcel freight logistics?

To reach this main question, the following sub-questions are proposed:

1. What kind of parcel delivery methods are there or will likely be developed in the coming years?
2. What factors influence the last-mile delivery choice of consumers and can consumers be grouped based on their preferences on those factors?
3. Which dynamic feedback loops quantify the connection between delivery services and user preferences?
4. What is an effective modelling approach to simulate the dynamics between consumer preferences and the delivery service?
5. To what extent do consumer preferences for last-mile delivery evolve over time under different service scenarios with their system capacity and price-setting?

1.5. Social & Scientific Relevance

As discussed in the previous sections, last-mile delivery needs to change, which is possible due to technological innovations. An increasing number of municipalities, like Rotterdam, try to influence shippers and couriers to perform a sustainable operation (*Convenant ZES*, 2020). Likewise, consumers will choose greener delivery options if clear information about their impact is provided (Ignat & Chankov, 2020). However, to make informed decisions, further knowledge about the dynamics in the last-mile environment is needed.

This study will provide a scientific contribution by developing a simulation model for the last-mile centred around consumer preference. This new approach can assess the dynamic interaction between consumers and last-mile operations qualitatively and quantitatively. The developed model will also integrate various delivery methods into one model, simulating their performance. Research with that kind of integration is limited, making it hard to predict operational performance in a quantitative way.

The social contribution of this study will be additional knowledge about the dynamics in the last-mile, which can help to develop sustainable delivery solutions. An example is that results could show in what kind of zones specific delivery methods are more desirable from a consumer perspective and an operational perspective.

1.6. Thesis Structure

In chapter 2, the last-mile environment and the current & future delivery methods are described. Subsequently, consumer preferences on those delivery methods and theories about preference evolution are discussed in chapter 3. The developed simulation model is presented first in chapter 4. Then in chapter 5, the simulation model is calibrated, verified and validated, after which the results are discussed. Lastly, in chapter 6, the conclusions and the recommendations for future research, model advancements and policymakers are discussed.

2

Current & Future Delivery Methods

This chapter will describe the last-mile environment. First, the existing delivery methods that are used primarily are discussed, and thereafter new concepts and innovations that are put forward in literature and by companies themselves will be presented. Although only van, self-collection, crowdshipping and drone delivery will be modelled, it is important to understand which delivery solutions could be available in the future.

2.1. Current Last-Mile Environment

Buldeo Rai et al. (2019) defines last-mile delivery as the final transport part of the supply chain from the last distribution centre, consolidation point or warehouse to the consumer. Thus the last-mile starts after long-haul transportation and ends when the parcel has successfully reached the consumer's preferred destination (Boysen et al., 2021). The literature review of *Revealing the secret emissions of e-commerce; Hint: it's all in the delivery* (2022) estimates that the average CO2 emission per parcel is 194 grams in Europe. The last-mile is further characterised by low-density transport, which is one of the reasons for the high costs, inefficiency and emissions that come with this logistic problem (Ignat & Chankov, 2020). In Figure 2.1, the forecast growth in delivery vehicles, emissions and congestion can be seen for the largest cities in the world when we continue in the present way. As the logistics network is reaching its capacity limit, delivery firms must create new delivery methods (Asdecker, 2020).

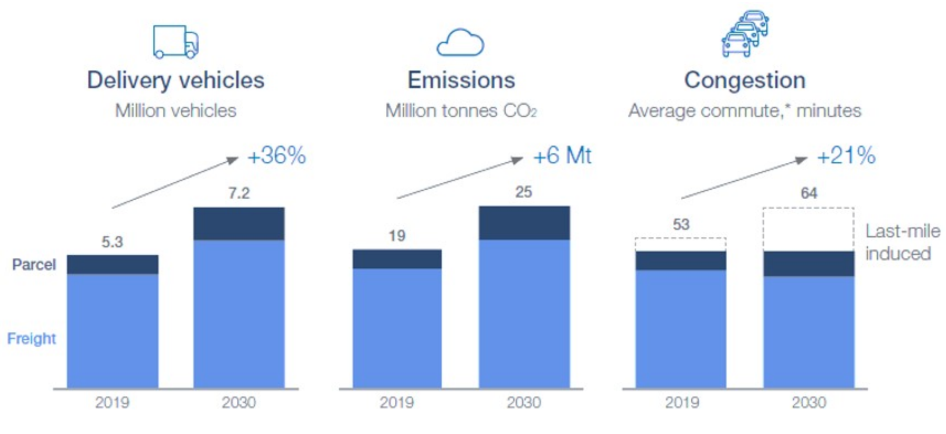


Figure 2.1: Impact of the growth of last-mile delivery according to Deloison et al. (2020) in the top 100 global cities. *Average commute for representative city.

Boysen et al. (2021) identifies five developments that pressurize the need for change: 1) *Increasing volume*, which is driven by urbanisation and e-commerce. Both urbanisation and e-commerce have seen fast growth in the last decade and will lead to a growing demand for parcel deliveries. Moreover, Cauwelier et al. (2023) observed that the online parcel order frequency shifted significantly from once in multiple months to monthly and weekly; 2) *Sustainability*, the growing demand brings more congestion and emissions, especially to inner-cities (Deloison et al., 2020). This results in an increasing focus on sustainable and environmental operations, fed by consumer awareness and governmental legislation; 3) *Costs*, traditional delivery methods come with high costs, and the commonly used attended home delivery (AHD) has the risks of failed delivery, 25% in the Netherlands (Buldeo Rai et al., 2019); 4) *Time pressure*, more and more delivery companies provide and more people expect fast delivery. Buldeo Rai et al. (2019) also indicates that these shorter delivery times complicate the possibility of efficient routing and consolidation; 5) *Ageing workforce*, working in the last-mile delivery business is physically demanding, whereas, in many western countries, the population is ageing. Creating a hiring problem for the required manpower of this growing segment.

Currently, two standard delivery options exist: AHD or collection via a locker or pick-up point (Buldeo Rai et al., 2019). The first option is undesirable from an environmental and logistical point of view due to the high rate of failed deliveries, less consolidation and low load factors of delivery vans, but it is the used delivery method for 82% of the parcels (ACM, 2020). Safe and secure Unattended Home Delivery (UHD) methods like personal reception boxes and in-car delivery are developed. Moreover, some companies try to nudge consumers into convenient delivery time frames from the company's perspective. Lockers and pick-up points have a 100% delivery rate and create the opportunity for efficient consolidation. Besides, consumers do not have to be at home for a successful delivery (Mancini & Gansterer, 2021). A disadvantage is that consumers are required to take the package to their home, which could also impact the transport network and impact the consumer experience. ACM (2020) found that only 16% to 52% of the Dutch people live within walking distance of a pick-up point or parcel locker, depending on the parcel carrier company. Most of these parcel points belong to one parcel carrier, but there is a development of 'white label' points that can be used by multiple companies. In the following part, currently used solutions will be described further.

Human-driven Delivery Vans

Currently, the universal delivery method for parcel delivery is a human driving a delivery van (Mohammad et al., 2023) (Boysen et al., 2021). The vans are loaded in a warehouse, consolidation centre or depot and deliver multiple parcels at different destinations in one tour. Tour formation is commonly done via a territory service; thus, each region is serviced in a repeated pattern. This simplification of the tour assignment comes with reduced flexibility and prevents optimising delivery tours on a daily base. Generally, parcels are loaded in a convenient manner so that they can be reached in order with the destinations. Adding multiple tours a day per region, for e.g. same day delivery, increases the complexity of the shipment-to-tour assignment problem.

Van delivery has multiple drawbacks that make it inefficient and disrupt other traffic. At each destination, the driver parks the van and goes door to door to hand over the parcel to the consumer, thus making this delivery method dependent on consumers being at home for a successful delivery and meanwhile potentially causing congestion. In the case of a failed delivery, there are three options: 1) Delivery to the neighbour; 2) Delivery at a pick-up point; 3) Return shipment to the warehouse. Sometimes, parcels are dropped off without directly handing them over to the consumer, but consumer consent is needed to do this. In addition to the previous problem, parking the van can be complex or disruptive. The research of Allen et al. (2018) concluded that the average delivery van stands still at 60% of the daily tour time and that a delivery person walks up to 12 km on foot. Lastly, this delivery method is impacted quite heavily by peak morning and evening traffic, extending the tour duration.

Cargo Bikes

A mode that is frequently used in cities is the cargo bike, often with electrical assistance (Boysen et al., 2021). Compared to vans, this mode can reach more restricted areas, and the parking problem is reduced, but the capacity is much lower. Therefore cargo bikes, in general, make multiple tours per day and are stationed at micro-depots. So the last-mile chain will usually be made up of a van transporting parcels from a depot to a micro-hub and the cargo bike transporting the parcels from the micro-hub to the consumer, which is called a two-echelon transportation system (Mohammad et al., 2023). Similar to the vans, this delivery method relies on consumers being at home for a successful delivery. The main contribution of this delivery solution to the previously stated problems is that it is a more sustainable solution than human-driven van delivery and that the impact on traffic is much smaller.

Pick-up Points

A regular and well-known solution to prevent failed delivery is pick-up points, which are shops that function as parcel point next to their ordinary business (ACM, 2020). This is an often used alternative in urban areas, and in heavily populated areas, it can be preferred by consumers (Demir et al., 2022). Compared to automated lockers, this option is more friendly for (older) non-tech consumers (Rózycki et al., 2021). Usually, pick-up points function as a collection and delivery point for consumers, and the study of Weltevreden (2008) concluded that these points are mainly used to return orders. This is confirmed by the low share of deliveries via pick-up points, only 5% in 2018 (ACM, 2020). Failed home deliveries are often brought to these locations (Rózycki et al., 2022).

(a) Delivery van (Kox, 2021)

(b) Cargo bike ("Coolblue start eigen bezorgservice per fiets in België: CoolblueFietst", 2018)

(c) Automated locker ("Beeldbank", n.d.)

Figure 2.2: Three currently operated delivery methods in the Netherlands

Automated Lockers

A quite similar and growing method for unattended delivery is automated lockers (Boysen et al., 2021), which are lockers that can function without the need for an intermediary at the location. These lockers are located in well-frequented areas and, in general, provide a 24/7 service, which makes them a convenient alternative for consumers that want to pick up the parcel at their time of choice. Additional benefits are that the carrier can fill the locker in off-peak hours when there is less traffic (Ranieri et al., 2018) and that a locker could function as a micro-hub (Mohammad et al., 2023). Micro-hubs are consolidation points in high-urban areas, where parcels are distributed to another delivery mode to perform the last yard of the delivery. For example, a carrier could drop multiple parcels in an automated locker. Subsequently, those parcels are collected and then delivered to the consumer by a cargo bike. The number of parcel lockers in the Netherlands is growing steadily but lagging behind in comparison with other European countries (ACM, 2020). In 2020 just 5% of the households in the Netherlands had a locker available within walking distance, and while according to Rózycki et al. (2022), it is essential for consumers that a parcel point is even within "slipper distance" (ca. 350m) in urban areas. Although delivery to lockers is theoretically an environmental last-mile solution, the work of Buijs and Niemeijer (2022) suggests that that effect is cancelled out easily by the additional travelling of consumers picking up their parcels. They indicate that a higher locker density can be more sustainable, as the likelihood of travelling by car reduces with shorter travel distances for consumers.

In Table 2.1, the comparison is made between a carrier delivering parcels to consumers' homes or a carrier delivering to a locker (Bilik, 2014). Self-collection points generally provide better vehicle routing and thus lower delivery costs (van Duin et al., 2020). Furthermore, drivers can work faster and more efficiently. However, a costly factor could be parcels occupying lockers for a long time. The potential cost reduction with mature self-collection is not consistent in literature. van Duin et al. (2020) predicts a potential reduction of 15% in the Pijp, a neighbourhood in Amsterdam. However, Seghezzi et al. (2022) estimated that operational costs could be up to 50% lower. Punakivi et al. (2001) even presented that self-collection can result in 60% lower operational costs.

Table 2.1: Comparison of carrier delivery and parcel locker delivery on a daily basis (Bilik, 2014)

Indicator	carrier	InPost parcel lockers
Daily kilometres / delivery driver	150	70
Parcels daily / delivery driver	60	600
C02 emissions / parcel	300 g	14 g
Fuel consumption / parcel	0.21 l	0.01 l

Electrification

Electrification itself is not a delivery solution. However, it can help to reduce the emissions that come with last-mile delivery (Deloison et al., 2020) (Buldeo Rai et al., 2021). Ordinary delivery vans can be fuelled by electricity or hydrogen. Also, hybridisation could be beneficial (Punte et al., 2019). Deloison et al. (2020) indicates that even in a conservative scenario, electric and hydrogen vehicles have the potential to reduce last-mile CO2 emissions by 16% and 24%, respectively. Table 2.2 shows that with electrification alone, the last-mile can still not be operated with minimal emissions, a combination with other innovations, like parcel lockers, is needed to optimise parcel transport.

Table 2.2: CO2 emissions per parcel for different adoption levels (Rózycki et al., 2021)

	EV adoption	**Parcels delivered by EVs in 2032:**		
Locker adoption		0%	50%	100%
Parcels delivered	0%	139 g CO2	86 g CO2	32 g CO2
to lockers	50%	87 g CO2	55 g CO2	23 g CO2
2032:	100%	36 g CO2	24 g CO2	13 g CO2

Thereby, the introduction of electric vehicles can only slightly reduce the operating cost of van delivery (Schröder et al., 2018). This is the consequence of the high share of labour costs in the total delivery costs, see Figure 2.3, especially in urban environments. Automated vehicles can significantly reduce operating costs because one operator can monitor multiple vehicles simultaneously.

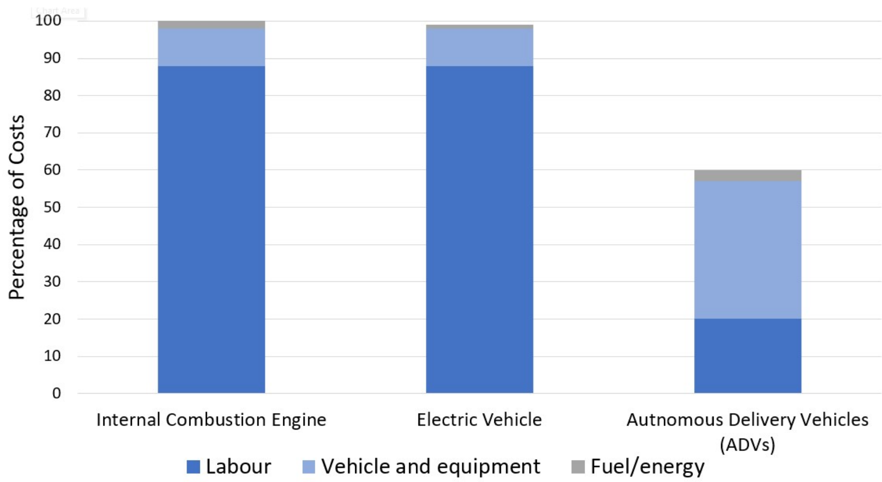


Figure 2.3: Last-mile delivery cost per parcel in an average city (Schröder et al., 2018). Assumptions: labour costs of €20/hour, average city-network density, and energy consumption of 0.3 kWh/km and 12.0 l/km.

2.2. Future Delivery Concepts

This section describes delivery concepts that are currently in development. Some concepts are already undergoing field tests, while others are still theoretical ideas.

Drones

A fast-developing delivery solution is drones, which are unmanned aerial vehicles (UAV) that perform AHD and UHD services (Asdecker, 2020) (Mohammad et al., 2023). Amazon, DHL and Alibaba are examples of companies currently doing real-world test cases (Boysen et al., 2021). Drones have the advantage that they do not depend upon dedicated infrastructure, which opens many possibilities in congested or rural areas. In addition, drones require less maintenance and come with lower labour costs than traditional delivery modes. These labour costs are lower as one operator can monitor multiple drones. In Sudbury and Hutchinson (2016), it is predicted that a delivery drone like the Amazon Prime Air will cost roughly 4.000 dollars. The operational delivery costs with drones are estimated at around €0.10 per kilometre (D'Andrea, 2014). Also, Aurambout et al. (2019) uses comparable costs to explore the viability of drones in European cities and Keeney (2015) estimates that Amazon could reduce their operational costs to $0.88 in the case of full drone network.

Safety and noise concerns must be tackled to become a fully operational delivery method, and range extension is asked. Thereby, drones have a small capacity, and many legal regulations must be overcome. Joerss et al. (2016) indicates that next to a low capacity, drones can handle a limited parcel weight. Furthermore, drones are currently quite big and thus require large landing areas. The combination of these strong and weak points implies that drones can be used for single, low-weight parcel transportation, giving drones mainly potential for rural delivery, especially for same-day and time-window deliveries. It is estimated that in Germany, this segment of delivery makes up 13% of the total deliveries (Joerss et al., 2016), indicating that drones could be a viable solution.

Another potential use case is launching drones from mobile platforms, such as delivery vans (Boysen et al., 2021) (Mohammad et al., 2023). This concept reduces the range problem of drones and consolidates parcels that drones can deliver in the same area. In even more future development, drones could also be launched from trains or flying warehouses.

(a) Delivery drone (Randall, 2019)

(b) Delivery droid ("Starship Technologies, Aramark launch contactless robot food-delivery service at ASU", 2020)

Figure 2.4: Examples of drone and droid delivery

Droids

Droids or unmanned ground vehicles (UGV) are comparable to drones but use the road infrastructure to move around (Boysen et al., 2021). The currently developed droids are intended for drive-walk use and thus travel at pedestrian speed. These slow speeds are compensated by the ability to transport much heavier parcels than drones. Next to that, droids can transport multiple parcels, thus combining different trips in one tour (Asdecker, 2020). Another advantage with respect to drones is that fewer legal barriers are expected for droids. Like drones, droids could be launched from a mobile platform like a van. It is expected that droids could be more reliable, safer and cheaper than human-supported delivery methods. Joerss et al. (2016) indicates that as long as the speed of droids is limited to around 5 km/h, cargo bikes seem to be a more cost-effective solution. If droids could also travel on ordinary roads, then their business case becomes competitive.

Crowdshipping

Crowdshipping is a last-mile solution that integrates passenger and freight transport by matching and connecting parcel consumers with travellers that have unused space or capacity to deliver the parcel (Gatta et al., 2018) (Punel et al., 2018). Goods are delivered by the crowd, which is provided with information via digital platforms and gets a payment per delivered parcel (Buldeo Rai et al., 2021). The crowd can use their everyday mode of travel; thus, parcel delivery could, for example, be done by car, bike or public transport. Crowdshippers can pick up parcels at service points, lockers or stores and then deliver them at another service point, locker or a consumers' home (Berendschot, 2021).

The advantages of crowdshipping are the flexibility, efficiency and scalability. Additionally, it opens up a labour group with fewer regulations and lower costs (Joerss et al., 2016). Punel et al. (2018) describe that due to the recruitment of crowdshippers via internet platforms, carriers get access to a large and heterogeneous group of people. At the same time, the costs can stay below a set limit. Berendschot (2021) assumes that crowdshipping is used when offered at a lower cost than other common delivery methods. In 2021 the average parcel delivery costs for B2C delivery was €3.65 (ACM, 2021), and thus crowdshipping costs will be below that limit.

The downsides are dependency on the willingness to act as a crowdshipper, safety concerns and inducing traffic. That last point refers to the rebound effect, where as a result of its own success, crowdshipping increases traffic and emissions as more passenger trips will be made to satisfy the freight demand (Tapia et al., 2023). To prevent this problem, for example, Gatta et al. (2019) looks at the potential of crowdshipping via a metro system because, in that case, the crowdshippers use ongoing metro trips. Joerss et al. (2016) indicate that they see crowdshipping as an excellent delivery method for market entrants due to the high scalability without the need for significant capital investments. Also, it can be a helpful addition during ultra-peak demand, such as Christmas.

Unattended Home Delivery

There are also developments for unattended home delivery, where the solutions overcome the problem of consumers not being at home and not creating the burden of picking up a parcel at another location (Boysen et al., 2021). Examples are reception boxes at the consumer's home, trunk delivery, and smart doors, which provide access to a certain room for a delivery person. These innovations do depend on consumers giving access to their private property. Certainly, trunk delivery and smart doors create privacy and theft concerns. These delivery options do not change the delivery process but make it more successful and, thus, more efficient.

Cargo Tunnels

Boysen et al. (2021) also mentions the use of tunnels for transportation of parcels in urban areas. UGVs or rail-bound vehicles could use these tunnels to move parcels from depots outside a city centre to micro-hubs in the inner city. From there, other modes could deliver the parcel to the consumer delivery location. The dedicated infrastructure comes with high investment costs.

2.3. Conclusions Last-Mile Delivery

To conclude, last-mile delivery is already a complex task, and the expected growth in demand means that new solutions are needed. Van delivery will come with high emissions, congestion forming and high costs. However, currently, it cannot be completely replaced by solutions such as pick-up points and cargo bikes. New innovations come with strong advantages, like high flexibility and lower operational costs. However, it is uncertain if innovative delivery methods, like drones and crowdshipping, will be adopted by consumers and if these new delivery methods can establish a well-performing operation. In the following chapter, the consumer perspective of the last-mile and the specific delivery methods will be discussed. In the later chapters, a simulation model will be presented that can represent the interaction between consumers and the operation of multiple last-mile delivery methods.

3

Consumer Preferences

This chapter will start with a discussion of attributes that are used to describe consumer preferences for last-mile delivery. Furthermore, it presents the consumer perceptions of delivery methods and descriptive groups of consumers found in earlier studies. This will provide the knowledge to simulate the initial preferences of consumers in the model. Secondly, it is explored how consumer preference evolves over time and which factors are of influence in this context. Lastly, the System Dynamics (SD) modelling approach will be discussed.

3.1. Attributes in Consumer Preferences

This section will begin with an analysis of the attributes that are used to describe consumer choice or preferences for parcel delivery in general. Thereafter, the consumer preferences for delivery methods are presented.

The acceptance of new innovations is in the hand of consumers, as research has shown that new products or services cannot be forced on consumers and that the impact of new technology relies on the readiness of the consumer (Vakulenko et al., 2019). In addition, Buldeo Rai et al. (2021) states that consumer acceptance depends on factors such as the service offer and service quality. For example, a study of over three-thousand respondents confirms that for 61% of the consumers, a positive delivery experience encourages them to shop with that retailer again. That makes it vital to understand which factors consumers view as important and how consumers look at different delivery methods.

A literature review is performed on attributes describing consumer parcel delivery choices and/or preferences. The literature selection is drawn from the Scopus database, and inclusion is based on strings and keywords. In Appendix A, the methodology used for the source gathering can be found. Table 3.1 presents the attributes applied in research, sorted on usage. Delivery speed is used in all papers and was generally divided into same-day delivery, tomorrow and 2-5 days delivery. The delivery costs are used in almost every paper, with one paper neglecting delivery costs as an attribute due to assuming free delivery. The costs, or fees, are generally offered in three ways: 1) free delivery; 2) free delivery from a threshold order amount; 3) a standard delivery fee. The possibility of choosing a time slot was used in five studies, with the option of time slots of two or four hours or no possibility for the consumer to indicate a time slot. Information services or parcel tracking came forward in half of the papers. Delivery reception was mentioned only three times, as not all studies accounted for delivery at different locations, like home or parcel lockers. Nguyen et al. (2019) and Buldeo Rai et al. (2019) used the attribute delivery date to explore the demand of delivery in the weekend. Two papers analysed how consumers deal with the knowledge on emissions by including an attribute for CO2 emissions. The remaining attributes were mentioned only once.

Buldeo Rai et al. (2019) presents that delivery price is the most important attribute for consumers. However, consumers are willing to wait for orders when the delivery time is more convenient. These findings are confirmed by Ignat and Chankov (2020), who add the insight that consumers are only will-

Table 3.1: Attributes for last-mile delivery in literature

Study	Delivery Speed	Delivery costs	Time slot	Parcel tracking	Delivery Reception	Delivery date	CO2-emissions	Daytime/evening	Return possibility	Delays	Particulate matter	Product Cost	Product Range	Carrier benefit
Gatta et al., 2018	X	X	X	X										
Nguyen et al., 2019	X	X	X			X		X						
de Oliveira et al., 2017	X	X		X	X									
Gatta et al., 2019	X	X	X	X										
Buldeo Rai et al., 2019	X	X	X		X	X			X					
Ignat and Chankov, 2020	X	X			X		X							X
Caspersen and Navrud, 2021	X			X			X			X	X			
Maltese et al., 2021	X	X	X									X	X	

ing to wait longer or pay more for greener delivery options if they receive clear information about the impact of their choice. Although Morganti et al. (2014) indicate that a German survey revealed that 29% of the consumers had negative experiences with delivery delays, only one paper used delay possibility as an attribute. In most studies, delivery speed is given as a deterministic option, where failed delivery or delays are neglected. Furthermore, Cauwelier et al. (2023) points out that the consumer delivery preference is independent of parcel weight and thus not relevant for preference consideration.

Results in Nguyen et al. (2019) show that delivery costs/fee has a far more significant influence than the other attributes. This indicates that consumers want to pay as little as possible for a delivery and that there is only a limited willingness to pay for faster delivery and shorter time slots. Furthermore, this paper distinguishes three different consumer groups, with segmentation that is firmly based on gender and income. Segmentation of consumers is likewise advised by Pani et al. (2020), as they conclude that there is considerable heterogeneity between consumers and that they cannot be expressed as a homogeneous group. Models can support the characterisation of different consumer groups and explain psychological, demographical and locational aspects influencing consumer preference. Even though there is variation across consumer groups, in general, trust, familiarity and lack of enthusiasm are major barriers for new technologies to be overcome. In the case of the last-mile existing and well-known delivery methods will co-exist with entirely new services. Vakulenko et al. (2019) indicate that when consumers are presented with a new delivery method, nearly all of them expect an improvement of the disadvantages of the previous delivery method, which influences the consumer expectations of the delivery innovation. But at the same time, consumers are more concerned with the drawbacks of a new innovation.

3.2. Preferences for Delivery Methods

The following part outlines consumers' points of view on the different delivery methods that will be modelled.

Van Delivery

Home delivery by vans is currently the dominant delivery method in the Netherlands (Molin et al., 2022). A major reason is that this delivery method is (often) offered for free. When choosing for home delivery, it is essential for consumers to be able to choose the delivery date and delivery time window, which is also as small as possible. That small time window increases the delivery success rates, but it has a negative effect on consolidation and routing efficiency (Rai et al., 2021).

Self-collection

The most critical attributes for self-collection points are the price, the quality and the location (Molin et al., 2022). The willingness to use this delivery method reduces with rising prices, with younger people being more price sensitive. The same study found a higher utility for increased delivery moments, which indicates that consumers value the flexibility lockers can offer with their extended opening hours. The study of de Oliveira et al. (2017) presented that consumers are willing to pay for a locker system, which the writers explain due to the high flexibility of this delivery method with respect to delivery time and location. As expected, the utility of self-collection points decreases with increasing travel distance of the consumer. The results of Molin et al. (2022) indicate that in a scenario where all delivery methods are free, which is the current standard in the Netherlands, 29% of the consumer will choose a self-collection point. This is a bit higher than a revealed preference study of the Authority for Consumers and Markets in 2018, which revealed that only 18% of the Dutch consumers did choose for delivery to a parcel point (Buijs & Niemeijer, 2022). If the consumer distance from the automated lockers is reduced from 2,5 km to 0,5 km and the pick-up point distance stays at 1 km, this percentage grows significantly to 57%. From this, it can be concluded that in the Dutch context, people are willing to use self-collection points if a dense network is provided. A considerable weak point of automated lockers is the lack of support, placing the burden of finding information or dealing with flaws in hardware or software at the consumer (Vakulenko et al., 2018). But the same study found that there is no clear consensus among consumers if the absence of human interaction during the service is beneficial or detrimental.

Crowdshipping

Although crowdshipping could be an interesting delivery solution, Buldeo Rai et al. (2021) describe expected issues with security, safety, privacy, reliability and accountability. They point out that the undefined character of the crowd is a source of stress. This is reflected by the low interest of consumers in crowdshipping found by that study with a survey of 1000 respondents. Only 21% is interested in receiving an order via the crowd, this percentage rises to 27% if the parcel is handled by someone living in the neighbourhood of the receiver.

A selling point of crowdshipping could be to offer the service at lower delivery costs than traditional van delivery (Punel et al., 2018). This is possible as crowdshipping theoretically comes with a lower operating cost. However, van delivery is commonly offered for free in the Netherlands, which makes this advantage not applicable in the context of this thesis. However, the study of Punel et al. (2018) indicates that current crowdship-users are dominantly persuaded by the eco-friendly characteristics of crowdshipping and not by the cost reduction. Another strong point of crowdshipping is the potential to offer a high level of personalization for the pick-up and delivery conditions. Gatta et al. (2018) explored consumer willingness to use crowdshipping and concluded that older people are less interested in using crowdshipping. The potential with younger consumers is confirmed by an exploratory study of Marcucci et al. (2017) on 200 university students, of which 93% indicated that they would accept to receive goods via crowdshipping. Additionally, Gatta et al. (2019) states that young people, men, and full-time employed individuals are most likely to use a crowdshipping service. Also, having a green attitude has a positive correlation with this service. That study presented a strong effect on the probability of adopting a crowdshipping service when there was no delivery date and time schedule flexibility. That probability was estimated to drop from roughly 60% to 12% when a flexible delivery schedule was not offered. Furthermore, crowdshipping has more potential in dense areas, with the chance of using crowdshipping increasing by 0.26% if the population density increases by 1% (Punel et al., 2018). Lastly, crowdship-users express that the system is not as easy to use as they expected in advance.

Another aspect of crowdshipping is that the crowd must be interested in shipping parcels. This supply side is labelled the Willingness to Work (WTW) by Tapia et al. (2023), and it is reported that socio-demographic factors have a considerable influence on the WTW as a crowdshipper. The most critical factors for accepting a delivery trip as a crowdshipper are the compensation and the maximum deviation time. This thesis focuses on the parcel receiver and will therefore not address this aspect in much detail, but it is taken into account in the MASS-GT version used in this study. However, based on this aspect of crowdshipping, this service is expected to perform best in an urban area (Punel et al., 2018).

Drone Delivery

The study of Merkert et al. (2022) suggests that consumers prefer conventional van delivery over drone delivery if all attributes are equal. On the other hand, drones can become very competitive when they fulfil the expectation that they are faster and cheaper. And currently, Amazon Prime and Wing offer 30 min delivery for less than US$1 in their test cases. In addition, another strong point of drones is their ability to offer a flexible service to various locations at all times (Leon et al., 2023).

An essential concern of consumers is the requirement of a quite large and safe place to land or drop a parcel (Merkert et al., 2022). Furthermore, the unmanned nature of drones can be regarded as unsafe and insecure. People indicate a fear of theft and of delivery to the wrong address or person. Additionally, the general public has privacy concerns with drones flying over and next to their property. The widespread operation of delivery drones seems dependent on robust evidence about the real-world performance of drone delivery services. Trust is pinpointed as one of the most crucial barriers to this innovation (Leon et al., 2023). Thereby, it is expected that consumers will trust a reputable company, like Amazon or PostNL, more than a new and unknown company.

Pani et al. (2020) reports that roughly 60% of the consumers are willing to pay extra to receive parcels using an autonomous delivery robot (ADR), but recommend an initial free or low fee for this new technology to strengthen its normalisation. Merkert et al. (2022) concluded that with a delivery fee of AU$2 a market share of 52% for drone delivery could be possible in a scenario with van and self-collection delivery. Even with high delivery costs AU$24, drone delivery could potentially gain at least 10%. However, the consumer preference for drone delivery depends on product characteristics. In Kim (2020), consumer choice probabilities for drone delivery are 7.18% for low-price clothing and 48.7% for urgent documents.

3.3. Last-Mile Delivery Consumer Groups

As expressed before, e-consumers are a heterogeneous group where preferences can be estimated based on socioeconomic and demographic data. A few studies performed a cluster analysis to describe the delivery preferences of different consumer groups. The scope of this study is not the analysis of heterogeneous consumer groups. Nevertheless, conclusions from other research can contribute to a better understanding of consumer preferences. Thus, the results cannot be directly used. However, some takeaways can still be noted.

Caspersen and Navrud (2021) performed a latent class analysis based on a stated preference survey with 513 respondents with a focus on the environmental attitudes of consumers and their behaviour. The study identified four distinctive consumer classes. The first class (31% of the sample) requests a high-level delivery service with environmentally sustainable solutions. They have a low willingness to trade higher CO2 emissions for reduced delivery time. The survey suggests that this class also plans their delivery well and believes that society should pay more attention to the environment. The class does not have distinctive socio-demographic characteristics. The second class (15% of the sample) has members with a higher income who use e-commerce more frequently. This class strongly believes in environmental change and is not time-sensitive. Consequently, they are willing to accept a longer delivery time if that decreases the environmental impact. Class-3 (41% of the sample) does not plan their delivery and is time sensitive; thus, they are willing to trade off environmental impact for faster delivery. This class does, again, not possess distinctive socio-demographic characteristics. The last class (13% of the sample) is made up of younger and more urban members. This class seems to care less about the delivery attributes than the other classes. They are not time-sensitive, 5-day delivery is preferred, and they are not really concerned about the environment. The study also provides an overview of some general remarks, see Table 3.2 that can be used when consumer segmentation cannot be realised or is complicated.

Buldeo Rai et al. (2021) performed a cluster analysis where 19.2% of the sample indicated a pronounced interest in crowdshipping. The study was focused on crowdshipping and based on a stated preference survey with 1000 respondents, which was described with a statistically optimal number of four classes. The ”trailblazers” (19.2%) were interested in crowdshipping and additionally, in the case of a failed delivery this class is also most interested in picking up a parcel at a neighbour’s house, while the other three classes prefer a pick-up at a parcel point. The results also suggest that this class is most open to other innovations within the last-mile. The analysis did not provide statistically significant socio-demographic characteristics for this class. Still, the authors indicate that it is probable that more

Table 3.2: Description of consumer types, their preferences and potential solutions, when little information is available on the consumer segment as presented in Caspersen and Navrud (2021).

Consumer type	**Preferences**	**Measures**
Age < 40	Time sensitive and prefer quick delivery, avoidance of delay and information about the shipment.	Unlikely trade increased delivery time to reduce emissions. Zero-emission vehicles with little change in the delivery solution might be an option.
Age > 40	Time is less important, although utility increases with reduced delivery time.	The least time-sensitive group of consumers who are likely to accept increased delivery time to reduce emissions.
Has income above average	Time is of average importance, with 5 days' delivery being preferred to one day. Emissions, especially particular matter, should be highlighted.	Likely to accept increased delivery time to reduce emissions. Potential measures might be consolidation or crowdshipping, which might take more time but have the potential to reduce emissions from last-mile transport.
Shop at least once a month	Consumers are both time-sensitive and negative to emissions.	Less likely to accept increased delivery time to reduce emissions. Technology investments that increase efficiency and reduce emissions might pay off if consumers return to the website.
Consumers believe in change for the environment	Less time-sensitive and accept longer delivery times for reduced emissions. Prefer information services.	Very likely to accept increased delivery time if the environmental aspect is highlighted. Expected to be flexible towards the delivery solutions: zero-emission vehicles like bicycles, crowdshipping, pick-up points etc.

frequent buyers have a higher likelihood of belonging to the ”trailblazers” class. A second class that is interested in innovations are the ”e-opportunists” (28.7% of the sample) and they are most interested in the flexibility of new solutions, but care less about the environmental benefits of last-mile solutions. Consequently, they are not willing to pay for a more sustainable delivery. Again, this segment did not possess distinctive socio-demographic characteristics.

The study of Pani et al. (2020) concludes that across all classes with rising age the willingness to pay (WTP) for UAV delivery decreases, which is explained by the lower technological efficacy and trust of older consumers. The study observed six classes from the data of a survey with 483 respondents with a focus on the use of e-commerce. In addition, it is found that males are slightly more open to delivery by autonomous vehicles than females. Furthermore, this study confirms that a respondent that has heard of a certain delivery method has a higher WTP, than consumers for which the delivery method is unknown. Lastly, the study indicates that the WTP for UAV delivery is higher in urban areas than in suburban and rural areas.

Nguyen et al. (2019) did research on the importance of different delivery attributes across multiple classes. A cluster analysis on convenience goods based on stated preference with 237 respondents provided three clusters. A ”price-oriented” cluster (36% of the sample) was found, for whom the delivery fee was by far the most important attribute. In the second cluster (22% of the sample), ”time- and convenience-oriented”, delivery fee is still the most essential attribute, but the other non-price attributes do in total account for 59% of the importance. Thus this cluster does value service attributes like delivery speed, time slot and delivery data. Lastly, there is the ”value-for-money-oriented” cluster (42% of the sample), where consumers seek to get the highest value for the money. Thus members of this cluster balance between clusters 1 and 2. Cluster one has a higher percentage of members with low income and females. While the opposite is applicable for clusters 2 and 3.

3.4. Evolution of Consumer Preferences

In this section, concepts that can be used to describe and model consumer preference evolution are discussed.

3.4.1. Perceived Service Quality

Perceived service quality is defined by Hepp (2018) as: "the gap between the perceptions of the actual service experienced and the expectations that the customer had before". The expectations comprise different factors, like past experience, personal needs, WoM and external communications. The received service consists of soft elements, like reliability and communication and hard elements, like the quality of the product/service and the proof of performance. Figure 3.1 shows how a consumer confirms its satisfaction by matching their expectations and their experience (Shabbir et al., 2017), in this case, called the perceived performance. This concept can be used to model the satisfaction of consumers and subsequently use that satisfaction to simulate a delivery preference at a future time point.

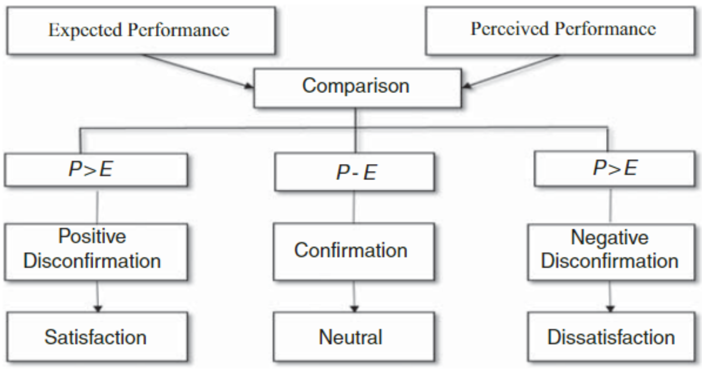


Figure 3.1: Service quality gap model from Shabbir et al. (2017)

3.4.2. Innovation Diffusion

Innovation diffusion theories or Diffusion of Innovation (DOI) explore technology acceptance outcomes based on a direct effect of a set of consumer beliefs, like perceived characteristics, compatibility, complexity, etc. (Wang et al., 2020). This theory stems from marketing, where it is used for the analysis and evaluation of life cycle dynamics (De La Torre et al., 2019). Additionally, it is applied for demand forecasting of new products.

According to Grawe innovation diffusion is the process of communicating an innovation throughout a network. Thereby the innovation diffusion consists of four elements: 1) innovation; 2) communication channels; 3) time; 4) social system. Thus when there is an innovation, information about that innovation is communicated via multiple channels across different individuals. Time represents the life cycle of an innovation, and the social system describes that individuals or organisations start working towards a common goal. Within the technology adoption life cycle, five sequential adoption groups are identified: innovators, early adopters, early majority, late majority and laggards (Chen et al., 2022). Those adoption groups, in general, have different demographic and psychological characteristics. An example is that innovators are commonly more educated and risk-taking.

The DOI theory is often complemented by attitude theories. DOI theory assumes a direct relationship between the perceived characteristics, like delivery speed and ease-of-use, of an innovation and the consumer adoption decision. While attitude theories suggest attitude is included as a mediator between beliefs and adoption intention; see Figure 3.2. As this study is more of exploratory nature, the DOI theory will be used as it requires fewer intermediate steps. Furthermore, this theory can be used to build up the SD model and hypothesise which causalities take effect, but it is not a direct simulation model or tool.

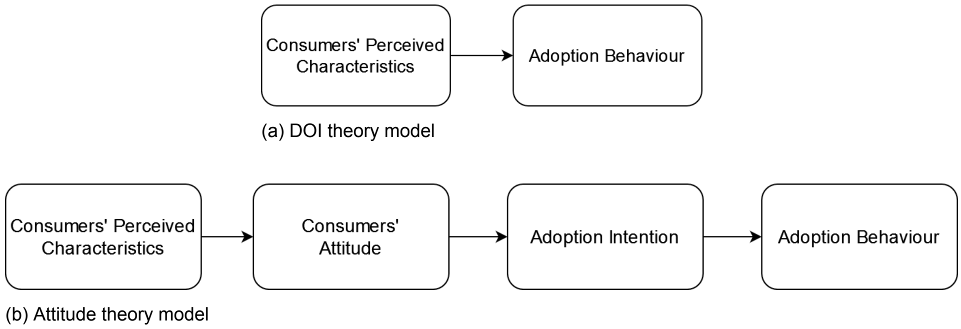


Figure 3.2: Difference between DOI and attitude theory

3.4.3. Word of Mouth

Word of mouth (WoM) is an important concept for studies that analyse the dynamics of user preferences and is defined by De La Torre et al. (2019) as: "the reduction of risks and uncertainty in customer acquisition and retention". It accounts for informal communication between consumers (C2C) that considers the characteristics and parameters of a business or a product. Thus, it describes the process of consumers evaluating a service or product and the purchase behaviour of fellow consumers and using that information for their own choice. It is reported that for adopting new businesses or services, consumers rely significantly on the experiences and opinions of other consumers in their decision process. WoM influences factors like satisfaction, loyalty, service level and trust. Furthermore, WoM is suitable when a population is heterogeneous or when the interactions between individuals are complex. Allsop et al. (2007) describes that WoM is best used in a holistic research program, which makes it a good fit with SD modelling.

WoM can take place in various forms, like neighbours talking to each other, e-mails, blogs and reviews. As not every consumer has the same social network, WoM spreads gradually and increases when more and more people have experience with a product or service, resulting in an 'S'-like adoption shape (Wong & Sheng, 2012). A potential mathematical representation for the adoption of a product due to WoM is presented by Wong and Sheng (2012) in Equation 3.1. Herein, k is a scalar constant that defines the shape of the distribution and α is a constant multiplier that describes how sensitive a consumer is to negative WoM.

$$Adoption = Consumers^{k(pWoM/(pWoM+\alpha*nWoM))} \tag{3.1}$$

Another and frequently cited formula for the adoption of new products or services by consumers is given by Bass (2004) and is called the Bass diffusion model. This formula describes that a consumer's initial purchase is related to the number of previous users and that sales of new products grow to a peak and then level to a value lower than that peak. It is theorised that an innovation is adopted by a group of innovators and a group of imitators. The first group makes their choice independently of others, while imitators base their choice on the social system around them. Over time all potential adopters are expected to purchase the service (*The Bass Diffusion Model*, 2023). Equation 3.2 provides the probability a new consumer adopts a product or service at time t if that consumer has not adopted it before, with t being the time in years. p is the coefficient of innovation and q is the coefficient of imitation. The average yearly value of p is 0.03 and generally ranges between 0.01 and 0.03. In the case of q, this is 0.38 with a range of 0.3 till 0.5 (Mahajan et al., 1995). m is the total number of initial purchases that can be made, and $Y(t)$ is cumulative adoptions till time t.

$$P(t) = \frac{f(t)}{1 - F(t)} = p + \frac{q}{m}Y(t) \tag{3.2}$$

The sales at time t can be found with Equation 3.3, where $P(t)$ is multiplied by the number of consumers that have not adopted yet. In the last formulation, the first part of the equation provides the innovation adopters and the second part the imitation adopters. This second part can thus be seen as the WoM effect (Mahajan et al., 1990).

$$
\begin{aligned}
S(t) &= P(t)[m - Y(t)] \\
&= pm + [q - p]Y(t) - \frac{q}{m}[Y(t)]^2 \\
&= p[m - Y(t)] + q\frac{Y(t)}{m}[m - Y(t)]
\end{aligned}
\tag{3.3}
$$

The Bass model is regularly used to analyse the market potential for innovative technologies and original diffusion patterns predicted by Bass have been shown to adhere with subsequent actual market outcomes (Massiani & Gohs, 2015). Thereby the Bass model has low data requirements. However, there are some disadvantages. Firstly, the model theorises that all potential adopters are eventually expected to purchase the new product. Hereby, only the initial purchase is considered and not repetitions. Secondly, the model is developed for new products or services, not for universally used services like van delivery. Furthermore, there is no input for other exploratory variables that could influence the diffusion, like prices or advertisements. Lastly, literature suggests that the effect of negative WoM should be explored (Mahajan et al., 1990).

3.4.4. Theory of Logistics Innovation

This theory describes that a firm's market share can increase because of a more effective logistics operation (Grawe, 2009). When firms identify a competitive advantage over another firm due to innovation, those firms will seek to establish the same innovation (Wang et al., 2020). In that way, innovations penetrate within that sector. Furthermore, logistics innovations often provide enhanced customer value, as firms can deliver solutions according to consumer needs at lower costs. Examples are the container and the steam engine. Various theoretical frameworks help to understand logistics innovation; three relevant frameworks are:

- The Schumpeterian innovation framework argues that large firms have a greater chance to innovate because they have a larger market power and larger R&D budgets. Thus the firm size and available resources determine firm innovation.
- The exploration-exploitation framework describes two different types of innovation. Exploratory innovations are designed for the needs of new markets, are radical and require new knowledge. Exploitative innovations target existing customers and are incrementally designed. This can be considered in the evolution of the different delivery methods, as, for example, drones are an undeveloped delivery method, while van delivery is a well-known and proven solution.
- The S-curve represents the evolution of radical innovations. The S-curve shows how consumer benefits increase slowly at the introduction of an innovation, that the benefits increase as the innovation develops and that the beneficial growth slows as the innovation matures.

3.4.5. Familiarity

Familiarity is defined by Fandos Herrera and Flavián Blanco (2011) as ”the number of product-related experiences that the consumer has accumulated”. It describes the understanding of a product and its characteristics by a consumer and the ability of the consumer to evaluate the quality of a product. A higher level of familiarity with a product tends to make consumers more trustworthy and loyal to that product. Vakulenko et al. (2019) point out that the introductory experience with a new service is essential for the future perception and expectation of that service. Thereby, the first use comes with unfamiliarity and a degree of confusion. However, consumers also self-educate over time, which can ease initial mistrust or concerns. In addition, when building up experience with e-commerce and delivery methods, consumers can reevaluate the importance of delivery attributes (Maltese et al., 2021). This is supported by the study of Johnson and Russo (1984), which explains that naive consumers find it harder to predict the performance of products, while consumers with high familiarity are able to select attributes that are predictive for the product performance. The familiarity theory can be used to model a kind of resistance or inertia of consumers to choose an unknown delivery method.

3.5. System Dynamics

An essential part of this thesis will be modelling the evolution of consumer preferences for delivery methods. The System Dynamics (SD) method will be used to achieve this. In this section, the choice for this method will be discussed and relevant earlier work that forms the basis for the conceptual model in chapter 4 will be described.

3.5.1. The System Dynamics Method

SD links qualitative and quantitative models via a causal loop approach and was developed in 1950-1960s by J.W. Forrester (Shepherd, 2014) and is based on the assumption that the behaviour of a system is mostly dependent on its own structure (Pruyt, 2013). Domains, where SD modelling is used are, for example, health policy, resources scarcity and supply chain management.

In dynamic systems, where time functions are relevant, SD can be used (Thaller et al., 2016). Hereby, a dynamic system is defined by Thaller et al. (2016) as ”a system where the external system environment causes observable changes in the internal system parameters”. The objective of SD is to explore the effect of decisions in a dynamically complex system, where SD supports in exploring and simulating non-linear feedback structures and functions. This is accomplished by connecting different system components and linking those connections with mathematical models. Angerhofer and Angelides (2000) identified that the essential viewpoint of SD is that feedback and delay cause the behaviour of systems.

SD describes a system at the aggregate level. It is a much-chosen representation of the internal process of an entity (Martin & Schlüter, 2015), like a parcel consumer or a group of parcel consumers. An advantage of this method is that it can visualise interdependencies between different parts of the system in a transparent way (Thaller et al., 2016). However, the multitude of relations can make it hard to grasp (de Almeide Correia, 2022). SD models can be used to perform medium- and long-term forecasts, trend analysis and impact assessments. A benefit of this method is the low data requirement, which is due to the high aggregated level of the model. Furthermore, it can integrate modelling algorithms from multiple fields, like economics and transport modelling. A drawback is that SD does not allow traffic assignment and point-in-time forecasts. The strengths and weaknesses described in research are summarised in Table 3.3.

Table 3.3: Strengths and weakness of system dynamics

Strength	**Weakness**
Visual communication	Complexity due to the number of relations
Representation of dynamic behaviour	Assumes homogeneity among agents
Can be qualitative and quantitative	Validation is complicated
Quantitative models can simulate the behaviour of a system over time	Equation based specification
Stability analysis	No point-in-time forecasts
Data requirements	No traffic assignment

An alternative is Agent-based modelling (ABM), as it is regularly used to study adaptive systems with self-organisation, emergence and adaptation (Martin & Schlüter, 2015). ABM helps to explain emergent patterns on a system level while considering the heterogeneity of entities, spatial and temporal heterogeneity, and stochasticity. But due to data scarcity, it is complicated to create a model that is grounded at the micro-level in this case.

SD structures

Rabe et al. (2020) illustrates that SD modelling is seen as a feedback process instead of a linear sequence of steps. Therefore a Causal Loop Diagram (CLD) is a useful way to describe the feedback loops within a system. Each feedback loop consists of at least two causality-related variables that connect back to themselves. Those connections can be positive or negative, and there could be a delay in the reaction.

The Stock and Flow Diagram (SFD) is defined as the underlying physical structure of the system, where the stock represents the state or condition of the system and the flow is changed by decisions due to conditions of the systems. This SFD determines the qualitative behaviour in the system based on differential equations that represent the feedback within the system. The variables or factors in the SFD can be categorised into four groups: level variables (stock), rate variables (flow), auxiliary variables and exogenous variables (De La Torre et al., 2019). A stock variable provides the state of the system at a point in time and is a variable that accumulates (Melkonyan et al., 2020). A flow variable changes the stock over time. An auxiliary variable presents information on stock and flow variables and is used to define intermediate concepts. And an exogenous variable is not part of the system's internal dynamics; thus, the system's behaviour does not affect this variable.

3.5.2. System Dynamics and the Last-Mile

SD has been used before to simulate the last-mile environment, additionally with a focus on consumer behaviour. Rabe et al. (2020) developed a model for automated parcel locker (APL) adoption within the last-mile environment. An SD model was combined with a facility location problem optimisation model to explore possible locations for APLs and their potential in the case of the city of Dortmund. In Figure 3.3 the CLD used for this study is shown. The potential number of e-consumers is positively influenced by the market size and the growth rate of e-shoppers. The potential of e-consumers does positively influence the amount of APL users. That number is also influenced by the APL market share and the APL growth rate, with the APL market share being an auxiliary variable fixed to 15% and the APL market growth rate being a flow rate. One of the positive feedback loops can be seen from the number of APL users to the number of APLs. The model showed, in the Dortmund case, that in 60 simulated months, the number of APLs did grow to cover the higher demand, and the geographic covering did also increase, although staying below the demand coverage percentage.

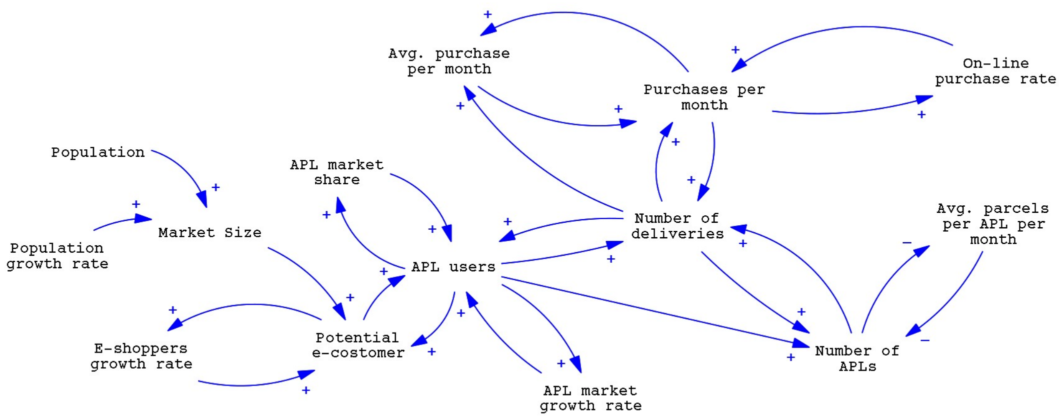


Figure 3.3: The CLD of Rabe et al. (2020) for the APL system

The study of De La Torre et al. (2019) performed an SD modelling technique to support the development of strategies and recommendations for the last-mile in the food industry. A dynamic modelling method was chosen because consumers adapt their consumption habits quickly, creating a dynamic environment. The development of new business models dependent on consumer behaviour is explored. The study identified four trends that affect the supply chain over time:

- Consumer consumption patterns: demand shifts
- Policy regulations: Internal and external standards to provide a high degree of supply safety with less environmental impact
- Decision-making processes: Focus on efficiency with minimal costs
- Technology use: Advances in technology made product tracking and improved delivery services possible

In the CLD shown in Figure 3.4 it can be seen that the variable new customers depends on multiple inputs. Firstly, advertisement increases the number of customers and a company advertises more when a higher gross profit is established. There is a reinforcing loop with the WoM variable and with leaving customers, adding to the group of potential customers. Additionally, the market size does positively influence the number of new customers. The attractiveness of the model provides a positive influence. This variable forms the connection with the operation of the service and describes the 'limits to growth' concept. This concept is used to model that a company's good performance will lead to a higher demand, which could reach beyond their capacities. When that happens, the company's offer becomes less attractive, and the model will respond. Lastly, the effect of collaboration is a positive influence on the number of new consumers. The study showed that providing a constant service level caused a saturation point and eventually a decline in customers, as people left the company and were not further attracted. Thus for a sustainable business model, an improving service level is needed.

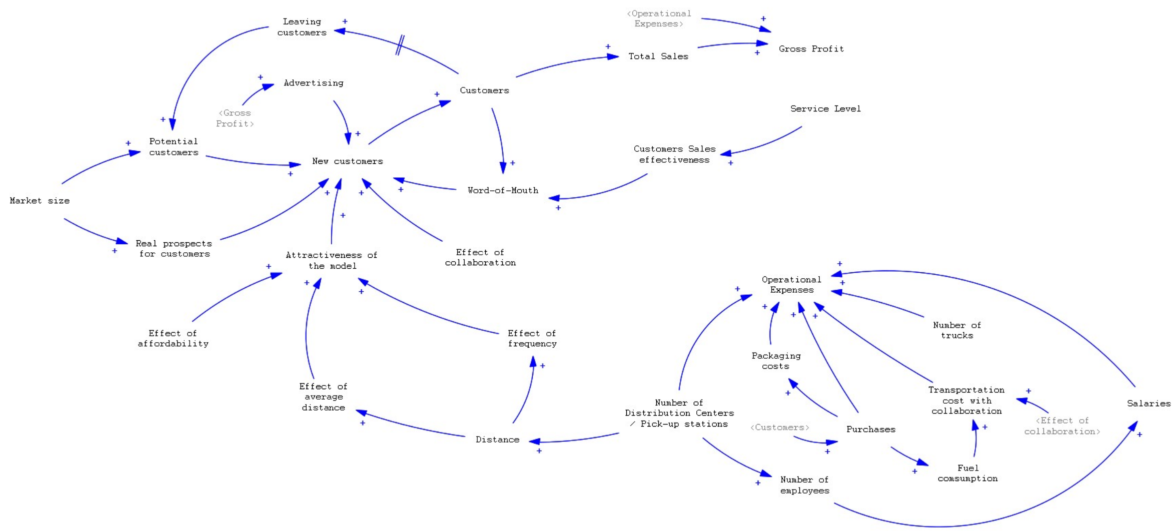


Figure 3.4: The CLD of De La Torre et al. (2019) for the consumer behaviour in the last-mile food industry

A third relevant study is Melkonyan et al. (2020), where three distribution channel options are compared for a local food and logistics provider, namely: 1) Centralised click & collect; 2) Decentralised home delivery; 3) Crowd logistics. The sustainability elements were quantified with an SD model, and a multi-criteria decision aid analysis was used to rank the most sustainable distribution options. In this CLD, the number of customers is impacted similarly as in the previously described models. New customers are reached by WoM, advertisement and relative attractiveness of the delivery model. Also, the market size does positively influence the number of customers. A higher service level positively influences customer sales effectiveness and, consequently, WoM. This model adapts the service level with a service improvement rate. The study concluded that the different distribution channel options provide a trade-off, with the crowd logistics option being the dominant option for almost all weight distributions in the multi-criteria decisions aid analysis.

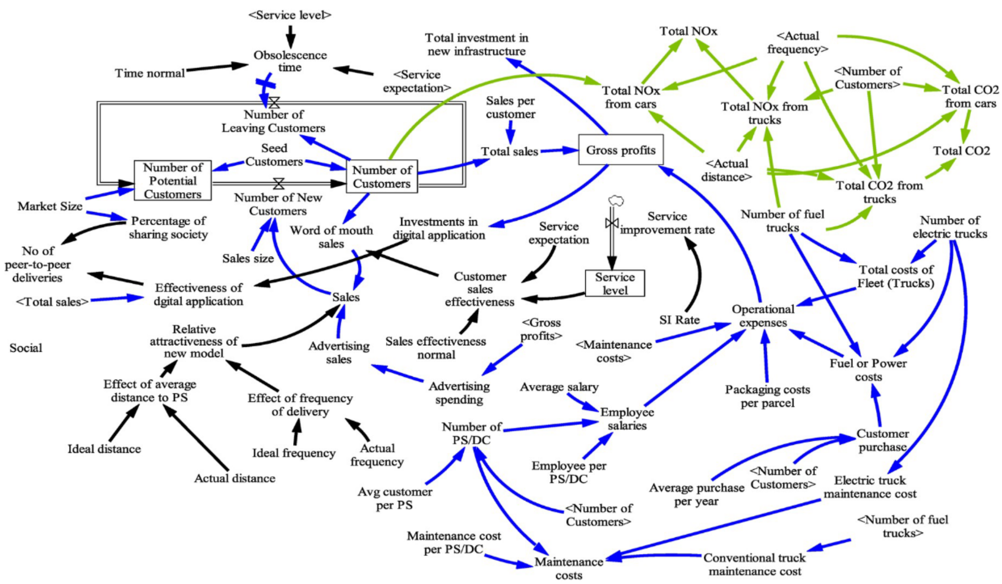


Figure 3.5: The SFD of Melkonyan et al. (2020) for a local food distribution network

3.6. Conclusion Consumer Preferences and System Dynamics

Multiple studies indicate that consumers change their preferences over time, also in the case of last-mile delivery. Due to causal effects, consumers evolve from an initial preference to an updated preference at various time points, and the supply side of the last-mile can evolve simultaneously. Innovation diffusion, Word of Mouth, familiarity and perceived service quality are theories that help to describe these phenomena. To integrate all effects, the System Dynamics method can be used because of its ability to represent dynamic behaviour, visualise the system structure and the low data requirements. Earlier research has shown that SD can be combined with other models to derive qualitative results and explore the last-mile environment and that the number of users is influenced by multiple variables within a system.

From the existing literature, it can be concluded that there is limited knowledge of how consumer preferences for distinctive competiting delivery methods evolve over time and how their relative performance influences consumer preferences. Present research only measures current preferences or uses a stated point estimate. It is expressed that more empirically grounded research is needed, and thereby it would create more insight into logistics if the consumer estimation models were connected to a traffic model. Furthermore, papers on new delivery methodologies often focus on technical or regulatory challenges, while knowledge of consumer acceptance is limited. Also, the contextual impact on delivery attributes for consumers is little explored. Finally, more research should be done on integrating multiple delivery methods in an estimation model for consumer preference.

4 Methodology

This chapter explains the structure of the developed consumer preference evolution model. First, the actors in the parcel last-mile are discussed. Secondly, the delivery methods that will be implemented are selected, and likewise, the attributes that are used to describe the delivery performance are selected. Thirdly, the conceptual model is presented with a general module structure and the used causal loop diagram. After that, a detailed description of the model parts is given. Lastly, the key performance indicators (KPI) that will be used to evaluate the model and the following results are discussed.

4.1. Actors

To conceptualise the last-mile, it is important to understand which actors are involved and which decisions they make. Roorda et al. (2010) point out that in the production and distribution of goods, multiple and diverse actors are involved, of which non has full control or knowledge about all the decisions. In general, three actors are specified: receivers, shippers and carriers. Quite often, one company fulfils the role of shipper and carrier.

In the context of this study, the receiver is the e-consumer. They initiate the demand for goods and services. Additionally, the receiver commonly determines the transport conditions, like the time or date of delivery (van Duin et al., 2012) or even the mode choice (McCabe et al., 2006). These conditions are stated when making an order at a shipper. The receiver regards the level of service as the most important concern when deciding on a shipper. As the carrier receives information from the shipper, the receiver and carrier only make contact at the delivery.

The shipper is a company that distributes goods to consumers and other firms (McCabe et al., 2006), with the last option not regarded in this study. The main decision for a shipper is the distribution of goods across their operating network, according to Ogden (1992). In the last-mile process, the shipper is responsible for product availability and the management and payment of the transportation of a good. As expressed earlier, it is assumed that products are always available in each depot of each carrier in this study. The exact time and route choice are usually not a concern for the shipper as long as the with the receiver agreed time or day of delivery is met. There are many shippers that own a private fleet and thus act as carriers. Otherwise, this part of the last-mile is outsourced.

Carriers are companies that move goods with their own vehicles, this can be different kinds of modes and intermodal combinations. A parcel carrier is often called a courier. McCabe et al. (2006) indicates that the main strategic decision for carriers is the amount and the types of vehicles they own, which is based on the service level they can provide and the corresponding profit and costs. Furthermore, the carrier decides how to move goods from their origin to their destination, thus performing the mode and route assignment. Those decisions could be influenced or dictated by the shipping actor. As these operating decisions are made for each shipment, the problem can become complicated if many shipments need to be combined. The carrier's vehicles are driven by a vehicle operator, which can be described as an additional actor. They can dynamically change planned routes and are responsible for the final delivery to the receiver. However, as this study does not model the last-mile in such detail, vehicle operators will not be considered as separate actors.

Additionally, government organisations are frequently discussed as influential actors in the last-mile environment (McCabe et al., 2006). They are responsible for all kinds of regulations and infrastructure decisions. This actor occasionally tries to adapt the status-quo or innovation directions by applying specific regulations or subsidies (van Duin et al., 2012).

4.2. Selection of Delivery Methods

As Boysen et al. (2021) indicates, a combination of delivery methods is needed because one delivery method alone will not overcome the five last-mile problems: Increasing volume, sustainability, costs, time pressure and the ageing workforce. Since user preference plays a critical role in adopting new delivery methods (Nguyen et al., 2019), this study will model several delivery methods and the consumer preferences for those delivery methods are simulated over time.

Considering that the focus of this study is on consumer preferences, the delivery methods that provide a distinctive service from the consumer perspective are regarded. For example, a consumer does not experience a significantly different delivery when a human hands over a parcel that is transported via a van or a cargo bike (Asdecker, 2020). This study will use van delivery, as that is the basic delivery method worldwide and can be offered in each case. pick-up points and automated lockers are quite similar, with the largest differences being a multi-purpose location for the pick-up point and longer opening hours for lockers (in general). It is chosen to simplify these options to one delivery method, "self-collection", which is a central location where consumers can pick up their parcel after delivery. This simplification is supported by the study of Molin et al. (2022), in which a stated preference survey revealed a relatively high correlation between pick-up points and automated lockers, which indicates consumers regard both options as quite similar. Related to the van/cargo bike argument, electrification is a valuable development but does not directly affect the consumer experience. It is, therefore, not included as a separate delivery option.

Table 4.1: Main strengths and weakness of the chosen delivery methods

Method	**Strength**	**Weakness**	**Parcel Delivery Costs**
Van	Capacity Accessibility Familiar method Usable for B2C and B2CC	Labour intensive Traffic dependent Parking requirement Emissions	€3.55 (ACM, 2021)
Self-collection	No failed delivery Consolidation Delivery and return	Consumer performs last-yard Location availability	€1.78 - 3.02 (Seghezzi et al., 2022) (van Duin et al., 2020)
Crowdshipping	Flexibility Scalability Efficient use of transport	Willingness to act as crowdshipper Inducing traffic Safety concerns	€1.50 - 3.55 (Berendschot, 2021)
Drone	Speed & instant delivery Road infrastructure independent Low labour costs No direct emissions	Limited capacity in volume and weight No-fly zones Space requirements for delivery Limited range	€1 (D'Andrea, 2014) (Aurambout et al., 2019)

Drones and droids have distinctive advantages and drawbacks and are currently implemented in small real-world test cases. Both delivery methods can potentially perform the deliveries that are currently done by a human-van delivery, but consumers will experience a different interaction due to the automation of these modes. Crowdshipping provides a unique delivery experience, as parcels are not handled by dedicated carrier personnel. This makes it sensitive to the consumers' willingness to use it. Although unattended home delivery can be really effective for the optimisation of last-mile logistics, it does not provide a specific delivery method. Lastly, cargo tunnels are not supposed for delivery at a consumer location but at depots and micro-hubs, thus making it an irrelevant innovation from the consumer perspective. Conclusively, this study will simulate van, self-collection, crowdshipping and drone delivery. In Table 4.1, the strengths and weaknesses of each delivery method are given, which shows that a combination of multiple delivery methods is needed to solve the last-mile problems.

4.3. Delivery Performance Attributes

As described in section 3.1, diverse attributes can be used to describe the performance of a delivery method, and with those attributes, consumer preferences can be forecast. The attributes that this study will use are:

- *Delivery Speed:* will provide an estimation of the delivery time of a parcel from the depot to the consumer delivery location. In this study, it is thus assumed that the product is available in the carrier depot, and prior transport is neglected. The levels are: several hours, same-day, 1-day or 2-day delivery.
- *Delivery Costs:* is a delivery method-specific fee with the levels: free delivery, cost below €2, below €5 and above €5.
- *Delivery Reliability:* will express the execution of the delivery by describing the way the delivery went. The attribute levels are: as expected, delayed, another delivery method, and unsuccessful delivery.
- *pick-up distance:* is similar to the delivery reception attribute and is used to model the distance a consumer needs to travel to receive their parcel. All delivery methods will hand over the parcel at the consumer's home, so this attribute is only relevant for the ”self-collection” method. In that case, this attribute provides the estimated distance a consumer needs to travel to the locker location. Those distances are then rated with 6 levels.

Time slots will not be varied over different delivery methods. Still, in the reliability attribute, the chance of a successful delivery for each delivery method will be mimicked, which accounts for the chance of people not being able to receive the parcel. And although order tracking is a regularly used attribute and seems to impact consumer's repurchase intention heavily (Nguyen et al., 2019), this study assumes similar tracking for each delivery mode. Also, emissions are not used as a specific attribute. This is acceptable, as Melkonyan et al. (2020) describes that consumers are rarely provided with information about sustainable product choices at the moment and that consumers with a sustainable mindset often do not show more sustainable behaviour patterns because of this.

4.4. Conceptual Model

To be able to simulate the evolution of consumer preferences and the operational fulfilment of the deliveries, a hybrid model is proposed that combines System Dynamics (SD) modelling and Agent-Based Modelling (ABM). The SD model will be used to model the preference evolution of consumers over time. The ABM simulates the carriers delivering the parcel from the depot to the receiver and the evolution of performance characteristics of each delivery method. In Figure 4.1, it can be seen how the hybrid model is connected.

The SD dynamics model and the ABM interact in two ways. One interaction is between the delivery operation, which follows from the carriers (agents) performing the parcel delivery, and the consumers evolving their preferences at the system level, resulting in a new demand for the carriers. The second interaction between the SD and AB models specifies that carriers can evolve their operation for

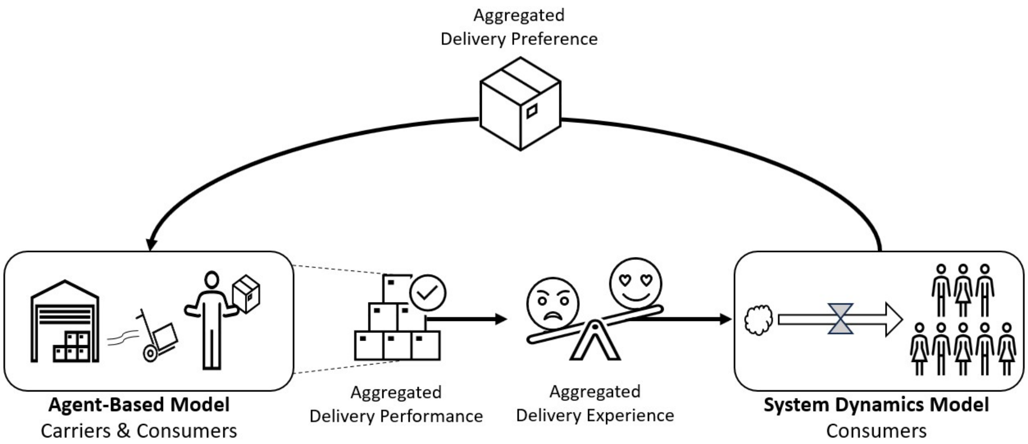


Figure 4.1: Schematic overview of the connection between the SD and ABM model

each delivery method based on the demand for that delivery method. With this structure, consecutive interactions can be simulated, and by that, data on consumer preferences and the delivery method operation can be obtained at multiple time points.

In this section, the general module structure is presented. Subsequently, the SD model is described with a causal loop diagram that simulates the relations between different factors in this consumer-centred, last-mile environment is discussed. Then MASS-GT, which is the used Agent-Based model. Lastly, it is elaborated on how the consecutive iterations represent time steps until the simulation horizon.

4.4.1. General Model Structure

Both the SD and AB models consist of various modules that are coupled in sequence. In Figure 4.2, the general model structure that shows how the modules interact with each other is presented. Each box represents a module, which is a model part that calculates a specific output. In green, the model parts that need to be developed are indicated, and in grey, the model parts that already exist. The modules within the light green plane will be part of the SD model, and the part in light yellow is the AB model.

Each simulation run will start with the zonal parcel demand, which is used as input for the initial delivery preferences of consumers. Those preferences and the parcel demand for each zone will be combined into a delivery preference for each parcel. In the zonal delivery demand module, the final delivery method will be assigned to each parcel based on preference and available capacity. That delivery demand will be grouped into depots that can fulfil the delivery, and eventually, delivery tours will be scheduled for each mode. Crowdshipping will be done via a separate module. The schedules can be used to estimate routes and corresponding emissions of the deliveries.

There are two feedback loops in this general model, indicated by the two arrows. The first loop describes the development of the delivery methods, which are based on the zonal delivery demand. The second loop describes the fulfilment of the delivery, where the performance of the delivery will be evaluated. The changes in the delivery method supply and the delivery fulfilment will then be used as input to develop new consumer delivery preferences, and a new iteration can be performed.

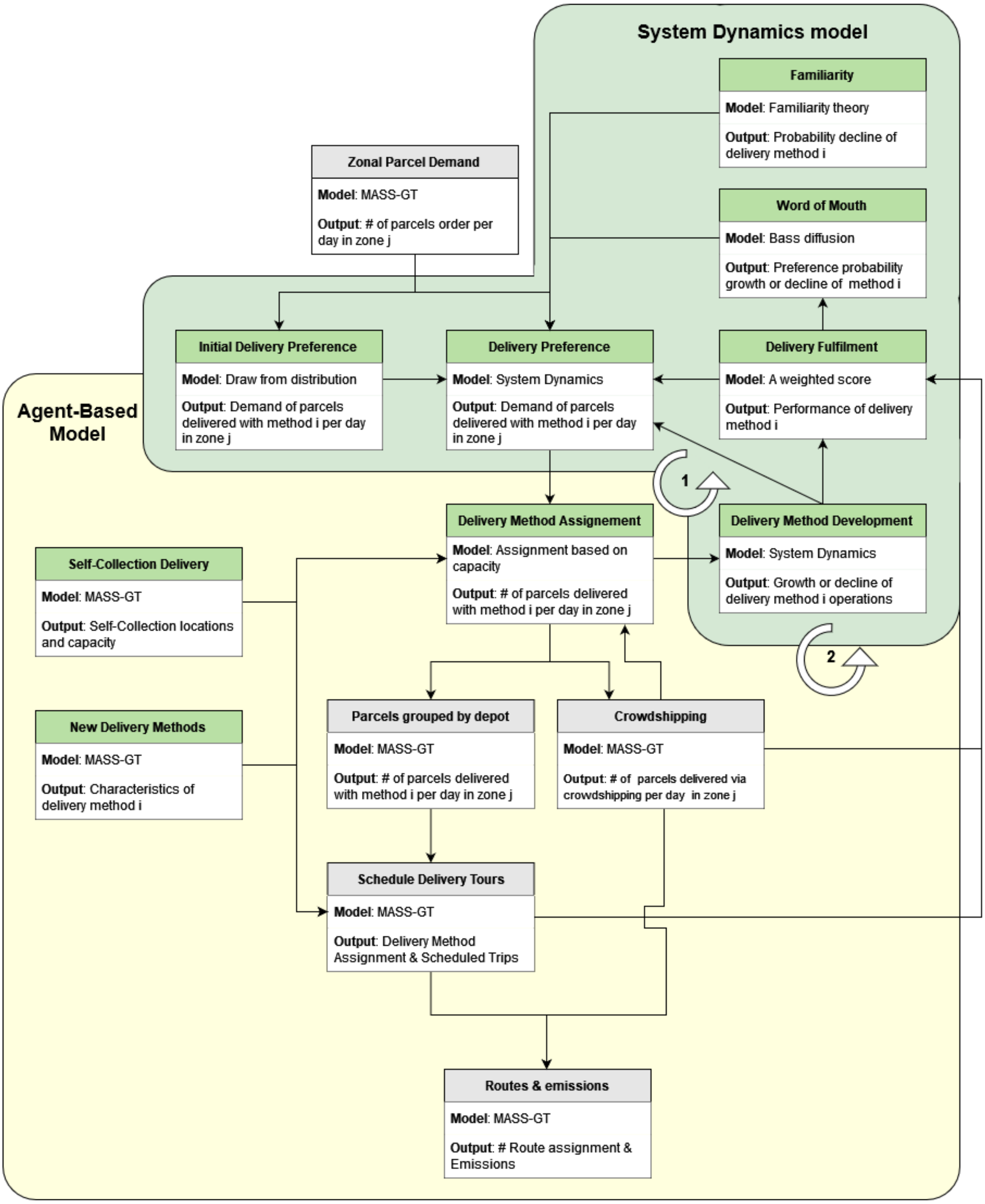


Figure 4.2: General Model Structure

4.4.2. Causal Loop Diagram

Both the evolution of consumer preferences and the operational supply of the delivery methods will be modelled with system dynamics. Based on the literature discussed in the previous chapters, the principal factors and their interactions are displayed in a causal loop diagram shown in Figure 4.3. All these factors, or variables, can be categorised as a stock, flow, auxiliary or exogenous variable, and in Appendix C, an overview of the variable characteristics can be found. Relations between variables are represented by an arrow, which is accompanied by a plus or minus sign. A plus sign indicates that a rise of the variable at the arrow tail also results in a rise of the variable at the arrowhead and vice versa for a minus sign. A delay is represented by two lines, ||, and commonly arises because time passes between a decision and its effects on the system (Sterman, 2002) or when time passes between the event occurrence and the information being known to the decision maker. Delays create the tendency of a system to oscillate. Feedback loops can be reinforcing (+) or balancing (-). In a reinforcing feedback loop, a change in one direction will constitute a change in the same direction. A balancing loop consists of a counter-reaction in the opposite direction. The SD model receives input from the AB model, and in a hexagon, it is indicated which components are ABM outputs.

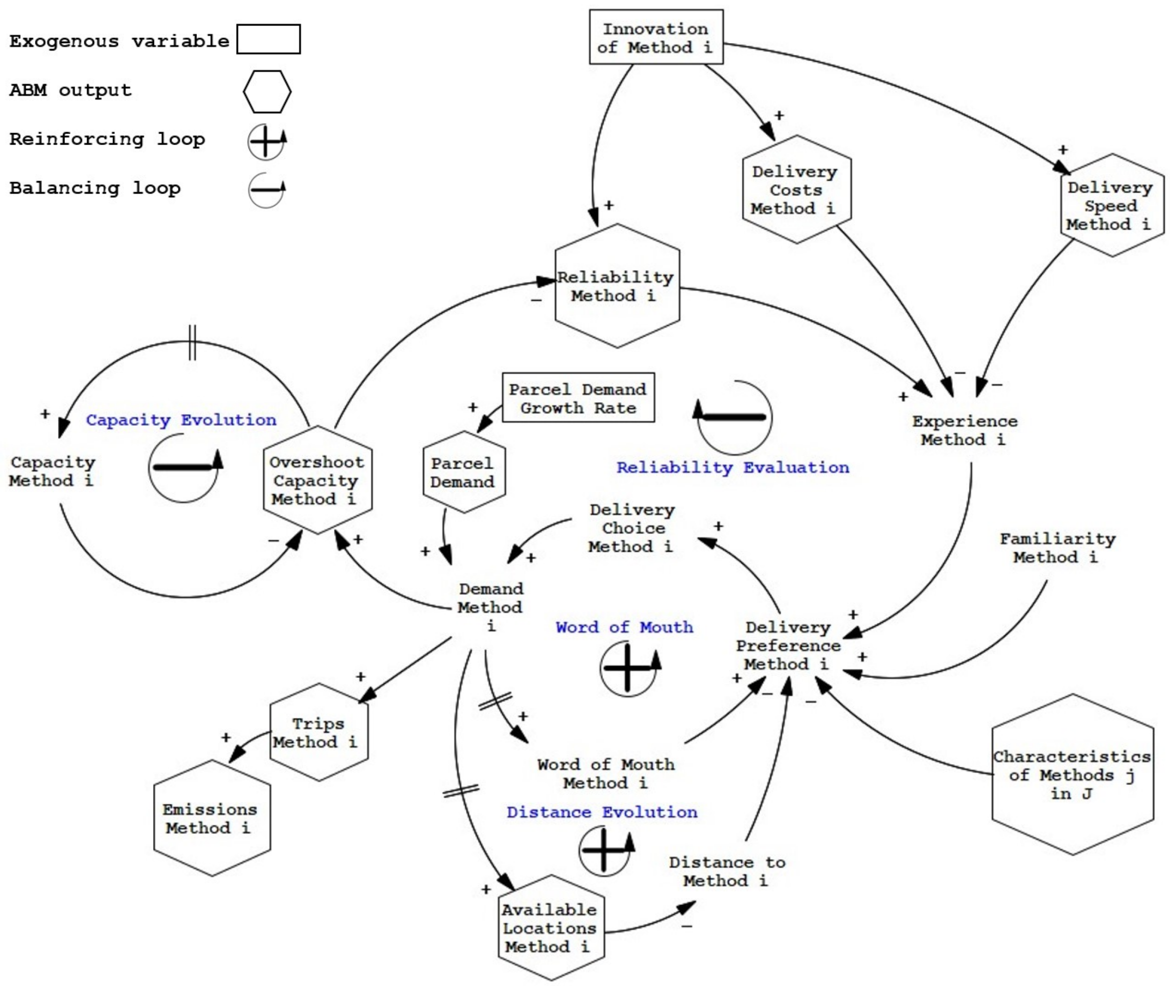


Figure 4.3: Causal Loop Diagram

The centre of the CLD is formed by *Delivery Preference Method i*, which leads via *Delivery Choice Method i* to the demand of each delivery method. As described in subsection 3.4.2, it is assumed, according to the DOI theory, that preferences directly affect people's behaviour. Furthermore, there are two kinds of influence factors: operational performance and mental mediation factors.

The operational performance will be expressed by reliability, costs and speed, specified per delivery method. Furthermore, the distance a consumer needs to travel to pick-up a parcel influences the operational performance. The factor *Reliability Method i* decreases as the capacity of a delivery method is overreached, as it is impossible to deliver those parcels. This reflects the 'limits to growth' concept discussed in subsection 3.5.2. When the demand is higher than the capacity, the capacity can be increased with delay, and vice versa when the demand is low. *Delivery Costs Method i* and *Delivery Speed Method i* are calculated per delivery; however, those factors are independent of the total demand. Considering that drones are a new technology, it is assumed that within the simulation time, this delivery method is improved due to innovations, which could enhance the operational performance.

It is expected that consumers also adapt their preferences due to non-experience-related factors. These are *WoM* and *Familiarity*. If those factors are positive, the preference for that delivery method will also be positively influenced. Finally, the demand for each delivery method will conclude in delivery trips, which causes emissions. As not all delivery methods will be influenced by the same variables, method-specific CLDs can be found in Appendix B. For all delivery methods, it can be seen that dynamic interactions will take place as there are multiple loops that reinforce (+) and balance (-) in combination with delays, which indicates that simulation is needed because of the system's complexity.

With the CLD, a Stock Flow diagram is established; see Figure 4.4. This represents the quantitative description of the variables. Stock variables are shown within a square, rate variables with ⋈, all other variables are auxiliary or exogenous. The main stock is the number of parcels that demands delivery via a specified method. This is fed by a flow of parcels per day for which a delivery method is chosen. *Capacity Method i* and *Available Locations Method i* are other stock variables that accumulate over time and provide a basis for the performance of a delivery method.

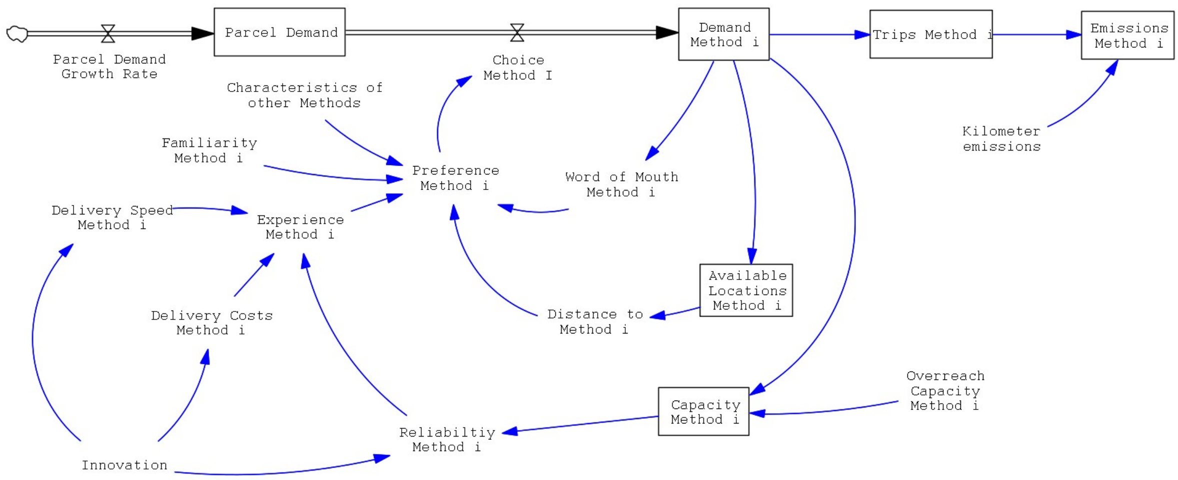


Figure 4.4: Stock Flow Diagram

4.4.3. MASS-GT

In this project, an existing agent-based model for urban freight transport can be used. The Multi-Agent Simulation System for Goods Transport (MASS-GT) is developed by Michiel de Bok and Lóri Tavasszy. It is still further developed in cooperation with Significance and the City of Rotterdam ("Mass-GT", 2023). MASS-GT consists of three modules (de Bok et al., 2022): 1) a shipment synthesizer that replicates the strategic processes behind freight transportation demand; 2) a tour formation model on the tactical level which assigns shipments to tours and vehicles; 3) a network model where route choices are simulated.

The model is calibrated on a number of data sets. One of them is a large data set form the CBS, wherein the Dutch truck trip travel diaries are collected. Information is gathered on the vehicle, the route and the transported shipments. This data collection was mandatory and automated, contributing to the large data set. In addition, another data set of the CBS on the firm population is used. From these data sources, the locations of logistic nodes were derived, and the shares of freight trips originating and arriving at these nodes were established. Thereby, the model accounts for multi-tier distribution. Thus, a parcel could travel via multiple logistical nodes, like a distribution and transhipment centre. This research will focus on the micro freight demand of parcels. MASS-GT models this parcel market by simulating the synthetic demand of parcels and then simulating the delivery schedules of delivery vehicles. During this study, not all modules and parts of MASS-GT will be used, as only parcel transport is simulated and exclusively the last-mile, thus only the transport from a depot to a consumer location.

4.4.4. Iteration Chart

The model will have a simulation time of five years. An iteration resembles a time step of a quarter of a year. Thus, the model will run until $t = 20$. The iteration chart presents which factors influence each iteration as some dynamics take place at different time steps. At the start point, the initial parcel demand and locations of the self-collection points will be set. In each iteration, the performance, WoM and familiarity will shape the consumer preferences. Each half year, the capacity of drones evolves based on the demand of the last two quarters. Each year the parcel demand grows with a constant growth factor, and the self-collection points evolve based on the demand of that year.

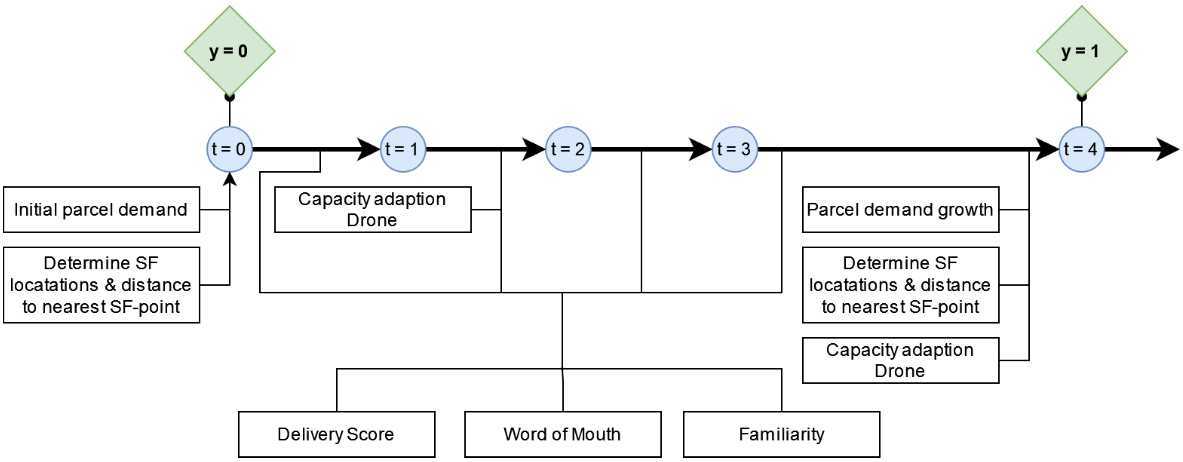


Figure 4.5: Iteration chart with the influential variables in one year

4.5. Model Specification

The general model, Figure 4.2, consists of multiple modules that are used to simulate the delivery of the parcels with different delivery methods to various locations. In this section, it is presented which modules are used and what their input and output are, hereby, the modules are grouped within the AB model and the SD model. First, the methodology to simulate each delivery method is elaborated.

4.5.1. Simulation of Delivery Methods

The four chosen delivery methods have distinctive characteristics for their operation. Van and crowd-shipping delivery were already implemented in MASS-GT. Self-collection and drone delivery are newly developed modules. Each delivery method is added in MASS-GT as an additional delivery method to which parcels can be assigned.

Van

Parcel delivery with vans is already modelled in MASS-GT. Vans start their tour at a courier depot, deliver the clustered parcels and end their tour at the same depot (Thoen, Tavasszy, et al., 2020). In the tour scheduling module, constraints such as capacity and tour duration are already considered. As the amount of vans per depot is not constrained, van delivery can handle the total demand.

Self-collection

In the current MASS-GT version, all parcels are delivered to the consumers' homes, self-collection at a pick-up point is included as a new delivery method. To include this delivery option, lockers and/or pick-up point locations are assigned to each zone or to specific locations. The distribution of the self-collection points is based on the Dutch average number of inhabitants per self-collection point. With that distribution, the pick-up distance for consumers to the nearest self-collection point is calculated. If a zone contains a self-collection point, it is assumed that it is placed at its centroid. In Equation 4.1, the formula to compute the average distance is shown. For each zone, the average Euclidean distance to that self-collection point is calculated by averaging the distance from all vertices of the zone to the centroid. With $D_{ij\ edge-centroid}$ being the distance from a corner point j of zone i to the centroid of zone i, J_i the number of vertices of zone i and $nLokcers_i$ the number of lockers in zone i. A uniform density is assumed over the entire zone, which provides the estimation that a consumer has to travel half of the average Euclidean distance from the vertices to a self-collection point in each zone. When there are multiple self-collection points in one zone it is assumed that they are placed across the zone and that the average distance reduces with each additional point. When zones do not contain a self-collection point, the distance to the centroid of the nearest zone with a self-collection point is calculated. The nearest zone is determined by the lowest skim distance from the zone without a self-collection point to zones with such a point. The skim distance is the shortest generalised distance across the modelled network in MASS-GT. The average distance is computed by averaging the distance from all vertices of that zone i to the centroid of the nearest self-collection point zone.

$$\mu Distance_i = \begin{cases} \dfrac{\left(\dfrac{\sum_{j=0}^{J_i} D_{ij\ centroid-edge}}{J_i}\right)}{(nLockers_i+1)} & \text{if } nLockers_i > 0 \\ \dfrac{\sum_{j=0}^{J_i} DnearestLock_{ij\ edge-centroid}}{J_i} & \text{if } nLockers_i = 0 \end{cases} \quad \forall \quad i \in Zones \tag{4.1}$$

Subsequently, for each request of delivery via self-collection, that parcel is assigned to the nearest zone with a self-collection point. Thus, in a zone with a self-collection point, parcels from neighbouring zones can also be delivered. For each zone with self-collection points, the total demand is calculated. If this demand exceeds the capacity, parcels will be randomly drawn until the capacity is reached. The remaining parcels are not delivered and will have a negative score for reliability. Based on the Amazon Hub each self-collection point has a capacity for 42 parcels per day (*Amazon Hub*, 2023).

Lastly, the assigned parcels for self-collection are added to the list of parcels for van delivery. This will constitute a van dropping multiple parcels at one stop. In that case, a drop-off time of 5 minutes is assumed per locker in each zone.

Crowdshipping
MASS-GT already comes with a module for crowdshipping for this study area in the version used for the HARMONY study (de Bok et al., 2021). This model assigns part of the parcels with an origin at a courier depot and a part that is local-to-local delivery. These are parcels that have an origin in one of the zones and are thus not limited to a depot location. Thereby, the assumption is made that all parcel requests can be picked up at a local store (Berendschot, 2021). For each crowdshipping request, the model allocates a crowdshipper that performs the delivery (Tapia et al., 2023). In that allocation, preferences from the sender are not taken into account, as they are assumed to be indifferent to the crowdshippers. The group of available crowdshippers is estimated based on the V-MRDH model (*Verkeersmodel geeft inzicht in verkeersstromen metropoolregio*, 2023). This specifies the origin, destination, travel mode and travel purpose of the potential crowdshippers.

Subsequently, the willingness to work as a crowdshipper is analysed for each parcel. This is simulated by a utility function for working as a crowdshipper (see Equation 4.2) and by a utility function for the regular trip (see Equation 4.3). Herein, remuneration is paid out to a crowdshipper to create an attractive utility for acting as a crowdshipper. The value of this remuneration is estimated with a natural logarithm between €1,50 and €3,35, with the last value being the average price for parcel delivery from B2C in 2020 (Berendschot, 2021).

$$U_{pickup} = \beta_{TravelCost} * (Cost - Remuneration) + \beta_{TravelTime} * Time + \eta_{pickup} \tag{4.2}$$

$$U_{currenttrip} = \beta_{TravelCost} * Cost_{trip} + \beta_{TravelTime} * Time_{trip} + \eta_{trip} \tag{4.3}$$

A commission of 15% is set to estimate the consumer costs of crowdshipping. That percentage reflects the commission across existing crowdship platforms like Nimber. With both utility functions, the most suitable parcel is allocated to the most suitable crowdshipper. The process ends when all parcels are allocated to the crowd or when crowdshippers with a higher utility for acting as a crowdshipper than for performing the regular trip are no longer available.

Drone
Drones have a capacity of one parcel, thus they repeatedly fly from a depot to a consumer and back to the same depot. As drones do not rely on road infrastructure, the flight distance is calculated by the Euclidean distance between the depot location and the centroid of the requested zone. A flight range of 10 kilometres is assumed (D'Andrea, 2014). Equation 4.4 shows the equation for the flight time of each drone delivery. V_{TT} represents the time that it takes to reach a safe flight height, the Amazon Prime Air flies at at least 100 meters Sudbury and Hutchinson, 2016, which is estimated at 60 seconds. D_{di} is the distance between depot d and the centroid of zone i, v is the average flight speed of a drone, which is roughly 45 km/h (D'Andrea, 2014), and a drop-off time of 120 seconds per parcel is assumed, which is also considered as the time to reload a drone at the depot.

$$DT = 2 * V_{TT} + 2 * \frac{D_{di}}{v} + 2 * DropTime \tag{4.4}$$

Drone deliveries are scheduled as follows: a delivery request is randomly drawn from the drone demand of a depot. The delivery time of that request is calculated and added to the total flight time of a drone. It is estimated that each drone operates 18 hours per day, from 6:00 to 24:00. For example, recharging and maintenance time is not considered. When the total flight time overshoots this time constraint, the remaining parcels will not be delivered via drone.

4.5.2. Zonal Parcel Demand

This component is part of MASS-GT and provides the number of parcel deliveries per day in each zone (Thoen et al., 2021), in this case only consumer parcels. The demand is estimated with an ordered logit model on an MPN survey with the explanatory variables: age, income and degree of urbanisation. The demand is a constant value and a yearly growth rate can be added to simulate the growing number of e-commerce orders over time. This module output is used as input for the Initial Delivery Preference in the first run and as input for the Delivery Preference in the subsequent runs. This module only provides the number of parcels and does not describe the weight or volume of a parcel. Thus the whole model will represent the number of parcels delivered by all methods.

4.5.3. ABM Modules

The modules in the AB modules estimate the operation of the deliveries. They are elaborated on in successive order.

Delivery Method Assignment

Here the demand for parcel deliveries by specific methods is converted into an assignment of delivery methods for each parcel. Not every parcel can or will be delivered with the preferred delivery method of the consumer. In Figure 4.6, a flow diagram is shown that represents the assignment procedure in this module. Each delivery method has an assigned speed score, as shown in the flow diagram, which is elaborated more in subsubsection 4.5.4.3. Firstly, for each parcel, the preference of the corresponding consumer is checked. If this is, for example, crowdshipping, the parcel will be given as input to the crowdshipping module. A feedback output of that module could be that a crowdshipper does not accept the parcel transportation request. In that case, the parcel is reassigned to van delivery. This comes with a delay penalty, a reliability score of 1 and a speed score that increases with two points. For all delivery methods, the demand should be within the (zonal) capacity. If this is not the case for self-collection delivery, this will imply that the parcel cannot be delivered, resulting in a reliability score of 2. If the capacity is overreached with drone delivery, the parcels are assigned to van delivery, with a delay penalty. For a successful delivery, consumers need to be at home when the crowdshipper, drone or postman is at its destination. A general successful delivery rate is used for all these delivery methods, which is 75% (Buldeo Rai et al., 2019). Of all parcels that are delivered via these methods, 25% will be randomly chosen as a failed delivery. The eventual output will be the number of parcel deliveries with a certain delivery method (i) per day in each zone.

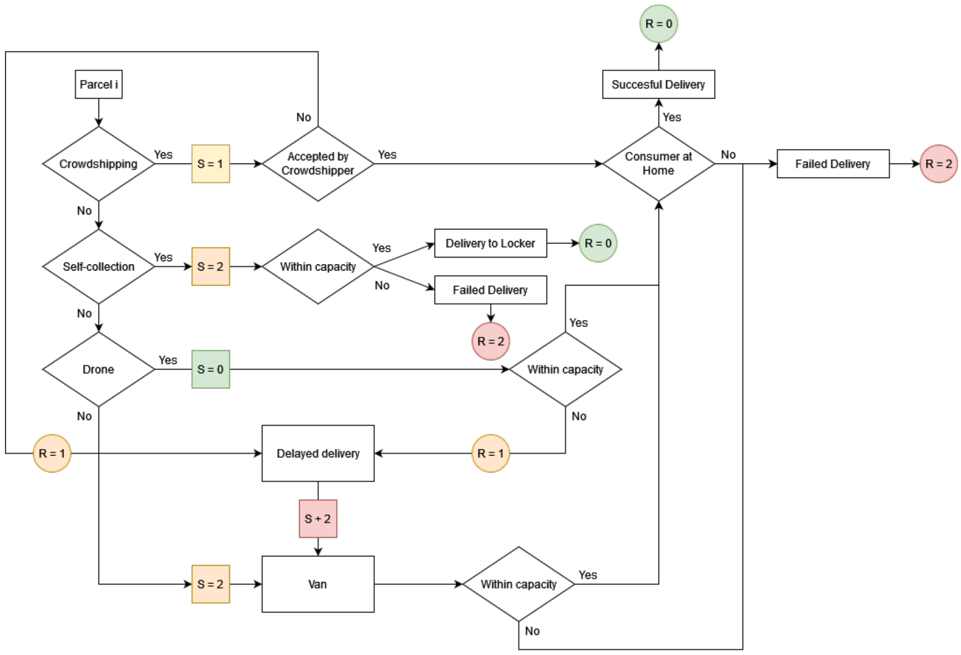


Figure 4.6: Delivery assignment flow diagram with Speed Score (S) and Reliability Score (R)

Parcels Grouped by Depot

This modelling step is part of MASS-GT, and no changes or additions are needed. The parcel courier for each parcel is determined based on observed market shared (Thoen et al., 2021). After the assignment to a parcel courier, the parcel is placed at the nearest, with respect to the delivery location, depot of that courier. It is assumed that the requested parcel is always available at each depot. Thus preceding transport is not modelled.

Schedule Delivery Tours

Again this is a component already existing in MASS-GT, where in this case, the new delivery locations and a new delivery method should be added as delivery options. The input is the assignment of each parcel to a specific delivery location and the mode of delivery. Then trips will be scheduled that combine the delivery of parcels with the same delivery method. Currently, MASS-GT groups parcels on geographical proximity and provides a plausible set of tours with constraints on trip duration and capacity (Thoen et al., 2021). For van delivery, parcels are grouped into clusters of 180 parcels maximum and with a maximum tour duration of eight hours. After the cluster forming delivery tours are scheduled, hereby, the nearest neighbour approach is used to minimise the tour distance. The time schedule is built up by the start time of the tour, which is randomly drawn from an average distribution across the day, a drop-off time of 120 seconds per unique delivery and the travel time between each stop. Drones will have a capacity of one parcel and will thus not combine multiple parcels in one tour.

Routes & Emissions
The following step in MASS-GT is determining the route for each scheduled trip. For van delivery, the scheduled tours are converted into trip matrices, which are assigned to a congested road network (de Bok et al., 2022). This assignment is based on an all-or-nothing shortest path assignment with generalised transportation costs of congested travel times. With those routes, the emissions for each used link are calculated using emission factors (g/km) (Thoen et al., 2021). MASS-GT makes a distinction in road type (urban, rural, highway) for these calculations and full and empty vehicles (Thoen, de Bok, & Tavasszy, 2020) with linear interpolation across the tour. Thus the output of this component are assigned van routes and the corresponding emissions.

The route calculations in MASS-GT estimate the route distance between the origin and destination zone. However, within each zone, multiple parcels can be delivered, for which additional kilometres must be driven; see Figure 4.7. The exact drop-off location for each parcel request is unknown. Therefore, the following assumptions are made to estimate the intrazonal vehicle kilometres and the corresponding emissions:

- In general, a uniform distribution of the population, and thus the parcel demand, across each zone is assumed.
- If the parcel demand of a zone is one, it is assumed that the request is located at the zonal centroid. And thus, no additional kilometres are added.
- When multiple deliveries take place in one zone for each parcel request, the additional distance is twice the average Manhattan distance from the centroid to the average Euclidean intrazonal distance. That average Euclidean distance is calculated by averaging the distance from all vertices of the zone to the centroid. And the average Manhattan distance for a point on a circle can be found by multiplying the Euclidean distance by 1.27 ($4/\pi$).
- For a self-collection delivery, it is assumed that a singular point is located at the zonal centroid. Because of that, when there is one self-collection point in a zone, no additional kilometres will be made.
- When there are multiple self-collection points and self-collection requests in a zone. It is also assumed that those self-collection points require intrazonal travel. That distance is in the same manner calculated as for van delivery.
- For the emission calculation, the intrazonal distance is multiplied by the emission factor for city roads.

Crowdshippers travel via the same road network as vans, and they make a detour from their original travel plans to pick-up and deliver a parcel. This detour distance is summed to gather the vehicle kilometres for crowdshipping. Furthermore, the MASS-GT model considers crowdshipping via car and bike. Because of that, the CO2 emissions are estimated by multiplying the emission factors (g/km) for cars times the detour distance of car crowdshipping. From 2013 the average CO2 emission of new passenger cars in the Netherlands dropped from 110 to 95.1 g/km in 2021 (*Average CO2 emissions from newly sold passenger cars in the Netherlands between 2008 to 2019*, 2023) (*Average CO2 emissions from new passenger cars, by EU country*, 2022), therefore an emission factor for car crowdshipping is set to 100 g/km. Drones perform simple tours, only flying from the depot to the consumer and back. Drones will operate on electricity and, therefore, not produce emissions.

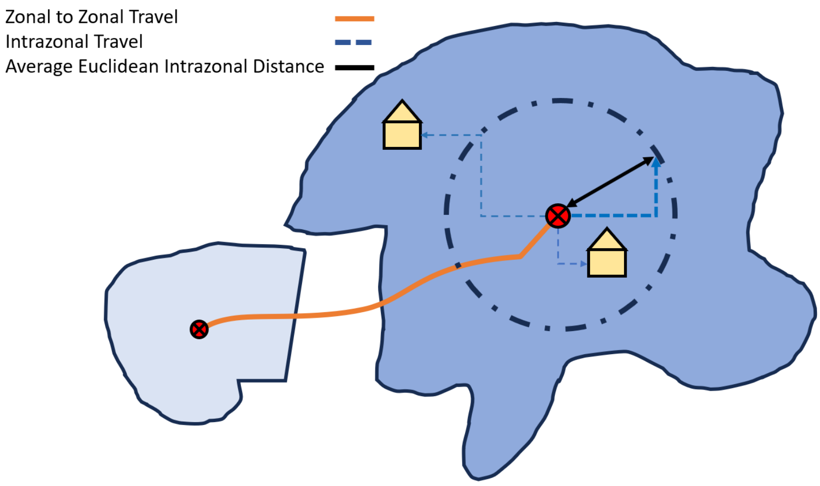


Figure 4.7: Distance calculations intrazonal trips

4.5.4. SD Modules

The SD modules represent the consumer preferences evolution and the development of the delivery methods.

Initial Delivery Preferences

The simulation in this study will start with initial market shares at an aggregate level for each zone in the study area. In Figure 4.8, the initial distribution of the preferences for the delivery methods is presented. The distribution is based on the assumption that consumers start with a preference for at-home delivery or delivery to a self-collection point. As described before, in 2018, only 18% of the Dutch consumers chose a self-collection point, while Molin et al. (2022) predicted that 29% of the consumers would choose this option, when both common and self-collection point delivery are offered free of charge. Based on this revealed preference from five years ago and the estimated potential, it is assumed that 20% of the consumers will opt for self-collection initially. Next, the remaining 80% of parcels are divided. Buldeo Rai et al. (2021) indicates that 21% of the consumers are interested in receiving a parcel via the crowd. As for drone delivery, significant steps in technology and regulations need to be made. Therefore, this delivery method will start with non of the consumers preferring drones. Drone delivery will develop during the simulation time, and its performance can convince consumers, thus adapting their preferences.

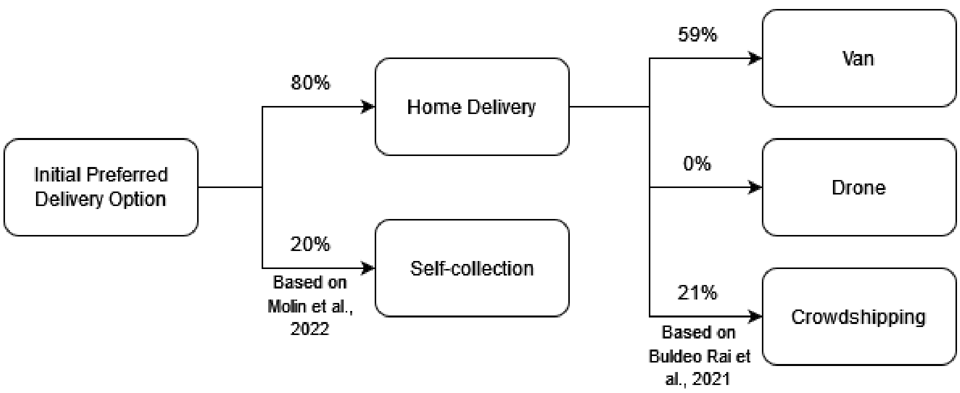


Figure 4.8: Division for initial consumer preference distribution

For each parcel, a delivery method is randomly drawn from the distribution as shown in Table 4.2. Thus this module provides the demand for delivery methods over the entire research area.

Table 4.2: Aggregate distribution of delivery method preference in the initial state

Delivery Method	Initial Preference Scenario 1	2	3
Van	80%	59%	59%
Self-collection	20%	20%	20%
Crowdshipping		21%	21%
Drone			0%

Delivery Preferences

This module will be the part where the consumer evolves its preference of the delivery methods based on the operational performance (experience) and due to word of mouth and familiarity. At the beginning of each run the preference distribution will be calculated based on inputs from other modules. The output will be given the same way as the Initial Consumer Preference and will overwrite this in each run. Then preferences will be drawn randomly from the distribution.

Delivery Fulfilment

To provide consumers feedback on their chosen delivery option, this component will estimate the characteristics of the performed delivery by calculating a performance score for each delivery method. This score can be seen as a quantitative representation of the perceived service quality of a consumer. Furthermore, it is theorised that according to innovation diffusion, these perceived characteristics directly influence the behaviour of consumers. As input for these score estimations output from the modules in subsection 4.5.1 are used. The score of each delivery method is calculated by:

$$S_i = \beta_s * \mu S_i + \beta_c * \mu C_i + \beta_r * \mu R_i + \beta_d * \mu D_i \quad \forall \quad i \in N \tag{4.5}$$

with μS_i being the average normalised delivery speed score of method i, μC_i the average normalised delivery cost score of method i, μR_i the average normalised reliability score of method i, μD_i the average normalised distance score for the consumer to pick-up their parcel with method i and $\beta_s, \beta_c, \beta_r, \beta_d$ being the weights of each attribute. The averages are computed with the attribute levels and normalised weights as shown in Table 4.3.

Table 4.3: Attribute levels

Attribute	Levels	Description	Normalised Performance Score
Delivery Speed	0	several hours	0
	1	same-day	0.04
	2	1-day	0.2
	3	2-days or more	1
Delivery Costs	0	free delivery	0
	1	costs < 2 euro	0.33
	2	costs < 5 euro	0.66
	3	costs > 5 euro	1
Delivery Reliability	0	as expected	0
	1	delayed delivery and/or other delivery method	0.33
	2	unsuccessful delivery	1
Pick-up Distance	0	distance < 100 meters	0
	1	distance < 300 meters	0.387
	2	distance < 500 meters	0.613
	3	distance < 1000 meters	0.774
	4	distance < 2000 meters	0.898
	5	distance > 2000 meters	1

Based on previous studies, the performance score of the speed attribute is exponentially growing, the score of costs is linear, and the pick-up distance grows in a logarithmic manner. The reliability attribute is assumed to have exponential growth. All attribute scores are normalised between 0 and 1. In section D.2, further explanation about the scoring of the attributes can be found.

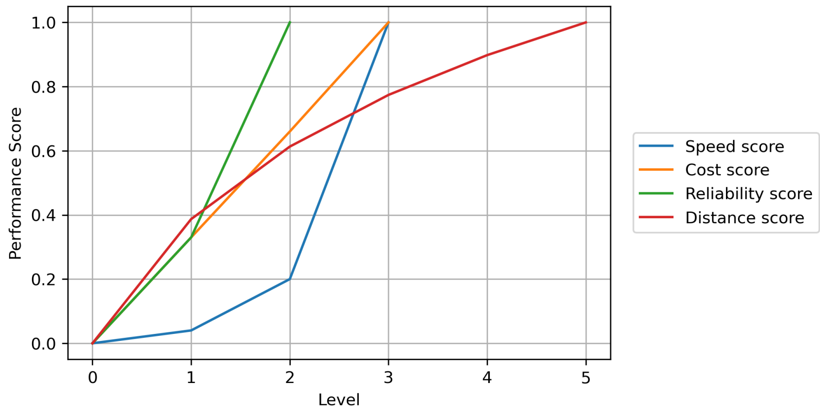


Figure 4.9: The normalised performance scores of the attributes

In Table 4.4, the initial levels of the delivery methods are presented. It is expected that drones can directly react to a request and thus deliver in several hours. Crowdshipping can use crowdshippers already travelling. Consequently, routes do not have to be scheduled, and it is estimated that parcels can be delivered on the day of ordering. Van and self-collection do currently provide a 1-day service in the Netherlands and are offered for free. The shorter delivery time of the other methods is expected to come with an additional fee. The pick-up distance is only applicable for self-collection delivery.

Table 4.4: The initial levels of the delivery methods with X indicating that a level score will be set after the initial run

Method	Speed	Costs	Reliability	Pick-up Distance
Van	2	0	X	-
Drone	0	1	X	-
Self-collection	2	0	X	X
Crowdshipping	1	1 or 2	X	-

To calculate the evolution of the delivery method probability over time, the performance of each delivery method must be linked with the probability of using that delivery method. A prominent approach for estimating discrete choices is a utility-based logit formula, with the Random Utility Model (RUM) the most generally applied version (Walker & Ben-Akiva, 2002). This theory describes that an individual derives utility by choosing an alternative, like a product or service. Thereby, it is theorised that each individual wants to opt for the alternative that provides them with the maximum utility. The utility is calculated with observable variables, in this case, the attributes delivery speed, costs reliability & pick-up distance, and an error term that accounts for unobserved variables, often called disturbances. With the utility of each alternative, the logit formulation provides an elegant way to estimate the probability of an individual choosing each delivery method. It is chosen to use a logit formulation as that is a well-known method to estimate preferences and choices. In section D.1, another formulation that was tested is discussed.

In Equation 4.6, the logit function used can be seen. There is no random error component included in this formulation. An important note is that the delivery method with the lowest μS score has provided the best service, therefore the performance score is subtracted from one.

$$P_i(t+1) = \frac{e^{1-S_{it}}}{e^{\sum_{i=0}^{N} 1-S_{it}}} \quad \forall \quad i \in N \tag{4.6}$$

Delivery Method Development

It is expected that certain delivery methods will grow or decline due to the evolving preferences of consumers. Based on previous demand, the service can be adapted, which will have consequences on the performance and ease of use of that delivery method in the future. Van delivery and crowdshipping will not evolve because van delivery has an unlimited supply, and crowdshipping depends on the availability of crowdshippers, which is assumed to be constant.

Drones are new, exploratory innovations. They require new knowledge to be developed, and it is likely that the benefit of their addition will follow an S-curve, as proposed in subsection 3.4.4 by the logistics innovation theory. Therefore the evolution of these delivery methods will be modelled with multiple aspects. The supply side can evolve based on the demand, the delivery speed can be decreased by product improvements, and the costs can be adapted following the adoption by consumers of the delivery method.

Drones are limited to a capacity of one parcel. However, depots can buy or sell drones based on the demand. This can be done on a six-month basis, thus every two iterations. If the difference between the demand and the offered capacity is at least 25% higher than the average number of parcels a drone delivers per day, an additional drone is bought by a depot. If this difference is negative and larger than 25%, a drone is sold.

To simulate the development of drones as new innovations, the product can be improved in this model, which can be expressed by reductions in the delivery time or large delivery area, which increases the daily capacity of drones. Three improvements can take place: 1) the vertical travel time is reduced; 2) the average flight speed is improved; 3) the range is increased. In Equation 4.7, the innovation probability, IP_j, taking place is presented. That probability is defined by the chance of a probability taking place P_I, set to 0.1% per run, times an increasing exponent based on the number of drones. This is done to represent that an innovation has a greater chance of improving if the innovation is highly adopted. If IP_j is true, the vertical travel time will reduce by 10% or the average flight speed or range is increased by 10%.

$$IP_j = P_I * 1.001^{nDrones} \quad \forall \quad j \in J \tag{4.7}$$

Lastly, the service quality of drone delivery can change by the costs of that delivery. The fee for drone delivery is determined as shown in Equation 4.8, with Pp being the profit margin of 15%, C_{drone} the average cost per drone per kilometre of €0.10 (D'Andrea, 2014) and D_{flight} the flight distance in meters.

$$Fee = \frac{((1 + Pp) * C_{drone} * D_{flight})}{1000} \tag{4.8}$$

Additionally, the distribution of self-collection points can evolve. With the assumption that building or removing happens on a yearly base, the self-collection points adapt every four iterations. If the demand for a self-collection point is lower than 75% of the capacity a point is removed from that zone. For that demand the total number of parcels delivered to that point is considered, as parcels for other zones could also be dropped at that point. A self-collection point can be added to a zone if the zonal demand is 50% higher than the capacity. The growth of the total number of points is limited to 20% per year. The priority for new self-collection points is based on the weight of the demand times the distance to a point. In that way, the points likely to improve the pick-up distance score most strongly are added.

Word of Mouth

To model WoM, consumers should evaluate their preferences based on information communicated via other consumers. The Bass diffusion model is used to do this, however, with one addition: the performance of a delivery method determines the effect of WoM, similar to Equation 4.9 proposed by Wong and Sheng, 2012 and Mahajan et al., 1990. The Bass formula is used to estimate the power of the WoM effect, but the relative performance of a delivery method dictates the magnitude and if that WoM is positive, leading to a large probability of preferring method i, or negative, resulting in shrinking that probability. Equation 4.9 shows the formula used in the model.

$$P_i(t+1) = P_i(t) + \frac{q\left(\frac{nParcels_i(t)}{\sum_{j=0}^{N} nParcels_i(t)}\right)\left(\left[\sum_{j=0}^{N} nParcels_i(t)\right] - nParcels_i(t)\right)\left(\frac{\sum_{i=0}^{N} S_{it}}{N} - S_{it}\right)}{\sum_{j=0}^{N} nParcels_i(t)} \tag{4.9}$$

$$\text{if } P_i(t) \neq 0 \quad \forall \quad i \in N$$

In this equation, q is the coefficient of imitation and it is set to 0.38, which is the advised average value as discussed in subsection 3.4.3. In red, the size of the population that can spread WoM about method i is determined. The blue part represents the number of consumers that did not use method i at time t. The part in teal provides the relative performance score of method i at time t. Thus, the part above the division line gives the number of new consumers, which is divided by the total number of parcels to obtain the additional probability of choosing method i. To maintain a summed total preference of 100%, the preference chance for each delivery method is divided by the sum of all preference chances, which is not shown in this formula to keep its structure clear.

Concluding, when a delivery method performs well and the relative amount of users is high, the probability of consumers preferring that delivery method in the next iteration grows. However, when most people have used that delivery method the WoM effect declines. The relative performance of a delivery method influences the magnitude and direction of the WoM effect. In Figure 4.10, it can be seen how the probability evolves over time for different performance scores.

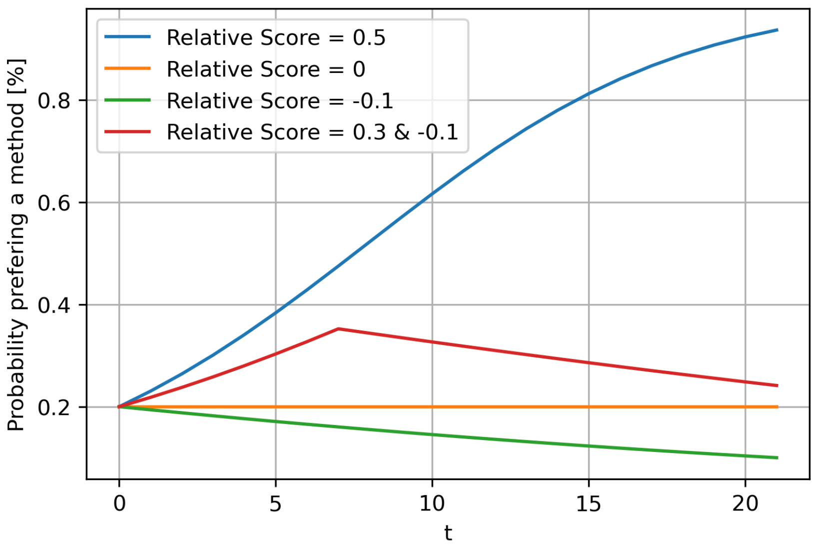


Figure 4.10: Influence of the Word of Mouth effect on the probability of preferring a specific method

Lastly, it could be possible that a delivery method is new and that based on the potential performance consumers will adopt it. Via WoM these consumers learn about a new delivery method and they will try it. The first part of the general Bass diffusion model is used to estimate the number of consumers that will test a new delivery method. As there are no previous users in this case the value of p can directly be taken as the probability that consumers prefer delivery method i. Thus when $P_i(t) = 0$ and the delivery score Sm_it is below the average delivery score, 3%, the percentage of innovators, p, in subsection 3.4.3, of the consumers will try the new delivery method at $t + 1$.

$$P_i(t+1) = 0.03 \quad \text{if } P_i(t) = 0 \quad \forall \quad i \in N \tag{4.10}$$

Familiarity
In other studies, familiarity is not directly used to estimate the probability of consumers adapting their preferences or choice. Consequently, a new methodology is developed based on the theory of familiarity. In principle, the familiarity theory describes that consumers build up trust and loyalty to a particular delivery method and that consumers hesitate a bit to use a delivery method with which they have no earlier experience. Therefore Equation 4.11 calculates the average chance that a consumer has never used a particular delivery method before. That probability is multiplied with Ω_{chance}, which represents the resistance to change and has a value of 5%. And likewise, the delivery score and WoM formula for the summed total preference should be kept at 100%.

$$P_i(t+1) = \frac{P_i(t) * \left(1 - \left(\prod_{P_i t=0}^{t} 1 - P_i(t)\right) * \Omega_{chance}\right)}{\sum_{i=0}^{N}\left[P_i(t) * \left(1 - \left(\prod_{P_i t=0}^{t} 1 - P_i(t)\right) * \Omega_{chance}\right)\right]} \quad \forall \quad i \in N \tag{4.11}$$

In Figure 4.11, the probability evolves over time for different starting probabilities and an Ω of 5% and 10%. Two main points can be concluded from this plot. First, the familiarity effect is stronger for an unfamiliar delivery method. Thus, the probability of consumers preferring an unknown delivery method does decrease relatively strongly, compared with a well-known delivery method. Secondly, over time the familiarity effect diminishes because theoretically, each consumer could have used that delivery method. For an unknown delivery method, it takes longer for the familiarity effect to converge to zero.

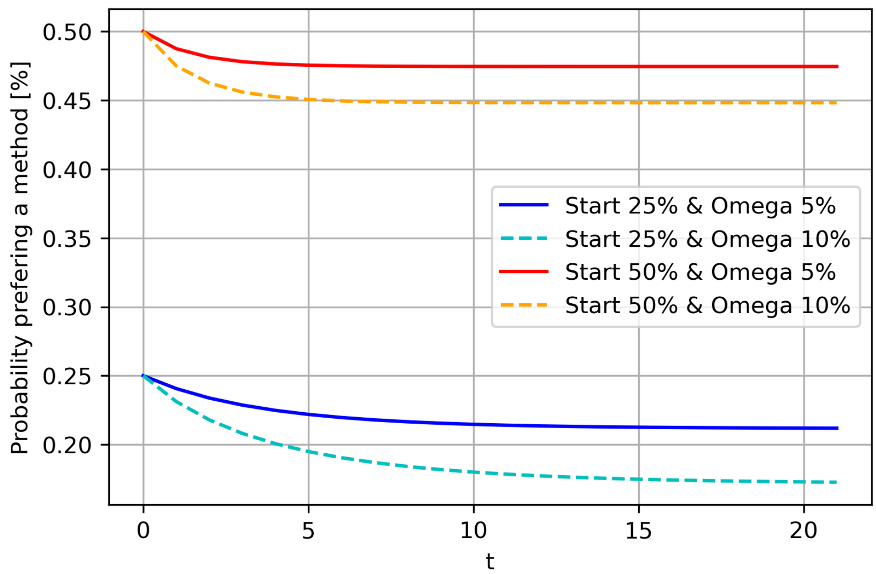


Figure 4.11: Influence of the familiarity effect on the probability of preferring a specific method

4.6. Key Performance Indicators

To evaluate the model results and the evolution over time, multiple KPIs are used to analyse the results. The main goal of this research is to analyse how the demand for delivery methods develops over time. Consequently, the KPIs should reflect the market shares of each delivery method and the number of parcels that are delivered. Additional information can be the amount of trips and emissions that are produced by the last-mile, as this provides insight into the benefit of the delivery methods from an environmental perspective. An overview of the KPIs and their characteristics can be found in Table 4.5.

Table 4.5: Overview of KPIs

Indicator	Method	Unit
Market share	Van Self-Collection Crowdshipping Drone	% % % %
Trips per day	Van Self-Collection Crowdshipping Drone	Trips/day Trips/day Trips/day Trips/day
Vehicle kilometers per day	Van Crowdshipping Drone	km/day km/day km/day
CO2 emissions per day	Van Crowdshipping Drone	kg/day kg/day kg/day
Number of Self-Collection points		Locations
Total capacity Self-Collection points		Parcels
Number of Drones		Drones

4.7. Concluding Remarks

In this chapter, it is discussed how the delivery fulfilment of each delivery method is modelled. Thereby, it is assumed that van delivery can provide a backup delivery service when crowdshipping or drone delivery cannot deliver that parcel. For both self-collection and drone delivery, a successful delivery rate of 100% is expected. In the cases of van and crowdshipping, this is 75%. None of the delivery methods has regions or zones where they cannot operate. Drones are modelled to fly a simple Euclidean distance from a depot to a consumer home, and that distance is used to estimate the delivery costs. The number of self-collection points is based on the average number of people per self-collection point, those points are then distributed based on the population of the zones. Consumers will evolve their preferences based on the performance scores of the delivery method, the WoM effect and the familiarity effect. Thereby, the level weights for the attributes are estimated on the results of earlier research. Lastly, the distribution of self-collection points and the number of drones per depot can adapt based on the demand. In the following chapter, the case study on which the developed model is applied is elaborated.

5

Model Application

This chapter discusses the application of the simulation model in a case study. First, the case is presented, and then the simulation scenarios are introduced. Secondly, the model is calibrated to improve the model's performance. At the end of this chapter, the results of the case study are discussed.

5.1. The Case Study

The developed simulation model will be applied in the case of the province of South Holland. South Holland is a highly urbanised region containing multiple large cities and the seaport of Rotterdam; Figure 5.1 shows a map of the area. The province has a population of 3.6 million, 50 municipalities and a surface area of 3.403 km^2. Both Rotterdam and the Hague are highly urbanized areas, yet, most other zones are still quite densely populated and cannot be called rural. South Holland has a well-developed transport network with various high-capacity roads and highways.

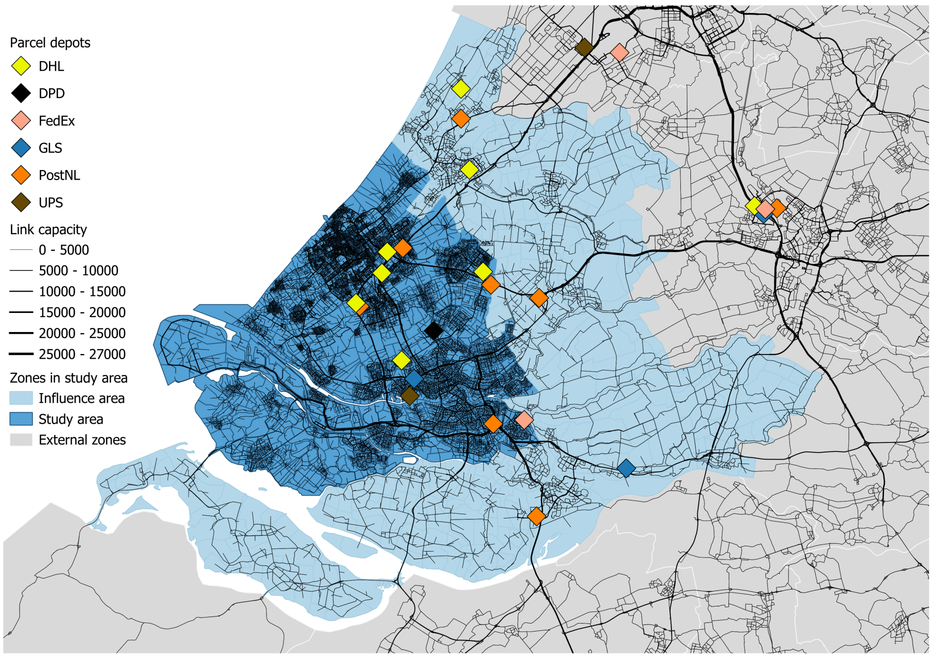


Figure 5.1: The study area of MASS-GT with the road network, zonal distinction and courier depot locations

As described in section 4.5, a parcel is assigned to a carrier based on observed market shares; in Table 5.1, the market shares in the Netherlands can be seen. Those carriers have in total 29 depots in and near South Holland.

Table 5.1: Carrier market shares (ACM, 2018)

Carrier	Share domestic	Share international
PostNL	62.5%	24.0%
DHL	27.5%	13.0%
DPD	2.5%	28.0%
GLS	2.5%	8.0%
UPS	2.5%	24.0%
FedEx	2.5%	3.0%

In MASS-GT, the study area is divided into almost 7000 zones of two types: study area and influence area. The 5925 study zones are located around and between Rotterdam and The Hague and are small with accurate road networks. The influence zones are less detailed. They are added as traffic in the study zones can originate, terminate or pass through the influence zones. In section 5.3, the presented results will describe the study zones.

5.1.1. Simulation Scenarios

With the presented model, the parcel freight logistics of South Holland will be simulated. This will be performed with a simulation horizon of five years. Each iteration inside a simulation run represents a time step of a quarter of a year. Thus each simulation will consist of an initial preference and 20 iterations to represent the evolution over time. subsection 4.4.4 shows the influential factors in each iteration. As the parcel demand grows each year, a constant yearly growth rate of 21.7% will be implemented based on Table 5.2. The population and its distribution will be considered constant. The following delivery scenarios will be simulated:

- Current State: only van and self-collection delivery. The self-collection points can evolve in number and capacity. A delivery success rate of 75% is assumed for vans and 100% for self-collection.
- Crowdshipping: As crowdshipping has to overcome fewer regulatory and technological barriers than drones, it is expected that crowdshipping will be the first delivery innovation that will be implemented in the Netherlands. This scenario will simulate the coexistence of van, self-collection and crowdshipping in the study area. The capacity of crowdshipping will be based on the availability of crowdshippers, which is kept at constant during the simulation. Again, the self-collection points can evolve in number and capacity, and a delivery success rate of 75% is assumed.
- Full innovation: Drones are added to the crowdshipping scenario. They evolve the number of vehicles and can evolve in performance due to R&D and innovation diffusion. As drones can make delivery within a smaller time frame, a success rate of 100% is assumed for these delivery methods.

Table 5.2: The yearly volume growth rate of parcel deliveries in the Netherlands according to ACM

Year	2017	2018	2019	2020	2021	Average
Growth rate	16%	20%	12.8%	34.8%	24.7%	21.6%

5.2. Model Calibration

This section will test and discuss the characteristics of the developed simulation model. First, a calibration of the self-collection points is performed to improve the consistency of the model. Thereafter the sensitivity of the model for the beta weights, the imitators coefficient in the WoM formula, the resistance coefficient in the familiarity formula, the innovation possibility for drones and different growth thresholds for self-collection and drone are elaborated. Finally, verification and validation tests are conducted to assess the model correctness of the model structure.

5.2.1. Calibration Self-Collection

Initial Distribution

In 2021 there were 10.698 parcel points, both dedicated postal stores and pick-up points, and 935 automated lockers in the Netherlands (ACM, 2021), thus a total of 11.633 self-collection points. Based on 17.800.000 Dutch people, there is one self-collection point per 1.530 inhabitants. It can be estimated that there are roughly between 2.300 and 2.400 self-collection points in South Holland, as the study area has a population of around 3.600.000 people. Based on the population of each municipality, self-collection points can be distributed across all municipalities, with, for example, the Hague having 340 points and Delft having 66. In zones with more than 1530 inhabitants, a self-collection point is added for each multitude of 1530. All other zones are assigned a self-collection point if the estimated maximum is not yet reached, with higher populated zones being assigned first. In Figure 5.2, the distribution of the self-collection points across the study area can be seen. It can be seen that mostly small zones, for example, the Hague or Rotterdam, with a high density, do have self-collection points.

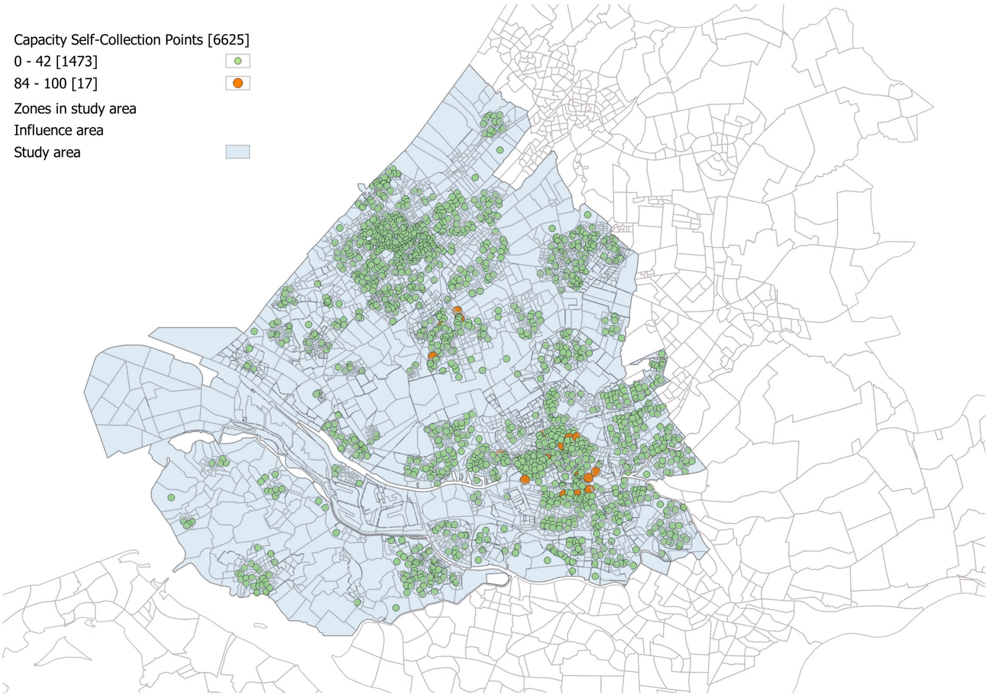
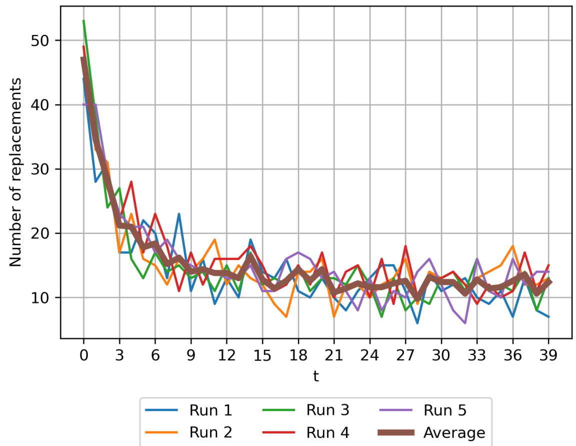


Figure 5.2: The distribution of self-collection points and their capacity over all zones at the initial stage

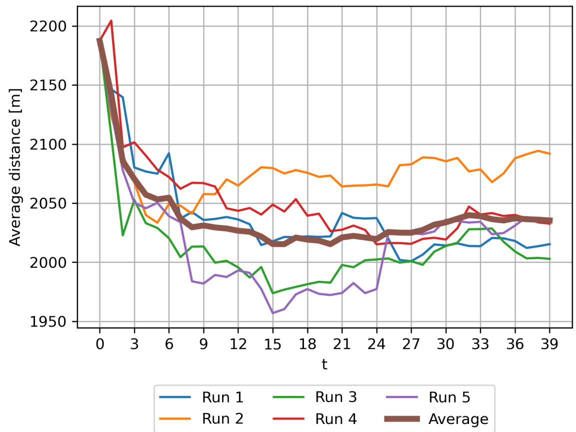
Calibration Process

Large fluctuations at the start of the simulation in the number and location of the self-collection points are undesirable, especially as the self-collection network in the Netherlands is quite stable. A calibration is performed to reduce effects from the self-collection evolution algorithm that redistributes the points to match demand and supply more closely. Consequently, after the calibration, the evolution of the self-collection points is dominantly an effect of consumer preferences. To establish a stable base situation, the distribution of the self-collection points is calibrated.

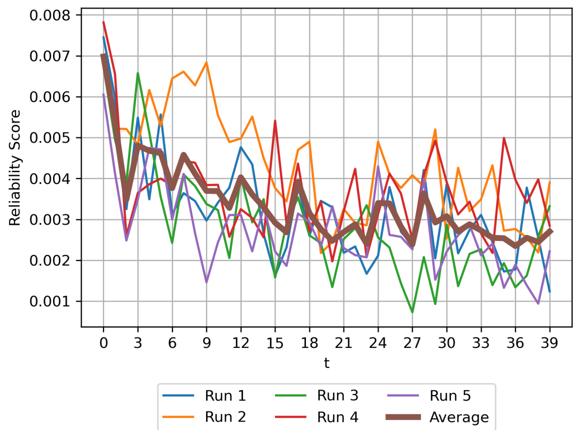
The distribution of the self-collection points is calibrated by iterating an evolution algorithm for allocating self-collection points multiple times. Hereby, the total number of parcels and the preference probability of self-collection per iteration are kept constant at 101594 parcels and 20%. In each run and each iteration, the parcel delivery method is randomly drawn from the preference probability distribution. Thus, a Monte Carlo simulation is performed. After each iteration, the number of self-collection points in each zone is adjusted in a different way, as explained in subsubsection 4.5.4.4. First of all, during the process, the total number of self-collection points in the study area is fixed. Thus, the removal of one point results in the addition of another point. For all zones, the demand for self-collection parcels is multiplied by the pick-up distance of that zone to determine an indication of the self-collection performance of that zone. If a zone without a self-collection point has a score three times larger than the score of a zone with a self-collection point, then that point is redistributed to the other zone. A margin of three is chosen to prevent displacement based on slight differences. This calibration process is performed five times, and the results are shown in Figure 5.3.



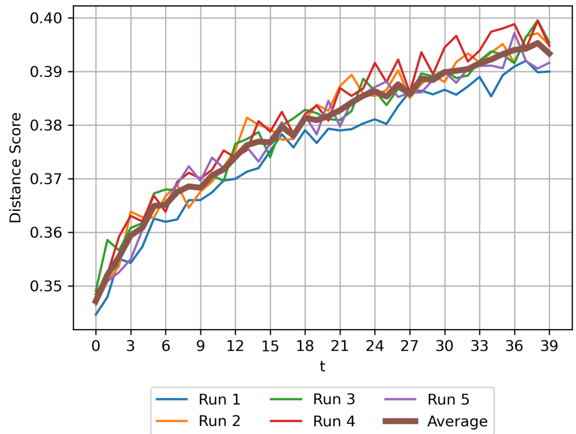
(a) The number of self-collection points that are redistributed



(b) The average distance [m] to a self-collection point for each zone



(c) The reliability score for self-collection

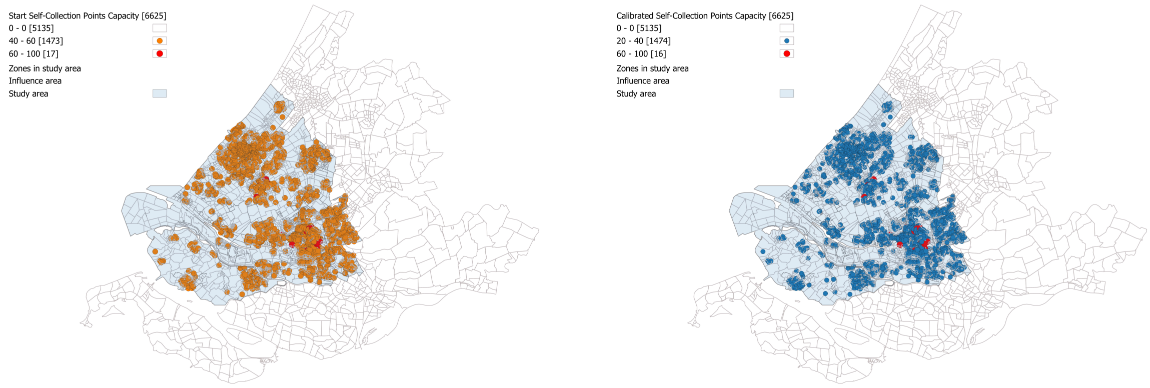


(d) The pick-up distance score for self-collection

Figure 5.3: The self-collection point calibration process for five runs

The results of the five calibration runs are largely similar. All runs start with 40 to 56 points that are replaced to another zone (a variance of $\sigma = 4.5$), which is followed by a fast reduction to around 12 changes per iteration ($\sigma \approx 2.3$ across iterations 9 till 39). Likewise, the average distance to a self-collection point shows a strong reduction that becomes relatively constant after ten iterations ($\sigma \approx 30.5$ across iterations 9 till 39). The reliability score also reduces, representing an improvement in the reliability, in the first iterations and converges to roughly 0.003 ($\sigma \approx 0.0009$ across iterations 9 till 39). During the calibration, the reliability score thus becomes twice as low. On the contrary, the distance score increases by approximately 15% during the calibration and seems to converge to a score of 0.4 ($\sigma \approx 0.0002$ across iterations 9 till 39). Lastly, the average demand per capacity of each zone is between 36% to 37%.

The redistribution of self-collection points becomes reasonably stable after 10 iterations. Therefore, the initial self-collection distribution is iterated 10 times to construct a base situation where the locations and capacities of these points are closer to the initial demand. As all runs are comparable, the self-collection point distribution of run 1 is used as the calibrated base situation. In Figure 5.4, the distribution before and after the calibration process can be seen.



(a) Before calibration

(b) After calibration

Figure 5.4: The distribution and number of self-collection points over the study area

Because the average load of the self-collection points is only around 37%, it can be concluded that the assumed capacity of 42 parcels per day is rather high. Therefore, the capacity of a self-collection point will be adjusted to 20. This makes it easier to get a picture of the evolution of the self-collection points. This adjustment is justifiable as this model does not account for parcels that occupy a self-collection point for longer periods of time or return parcels that take up capacity. Lastly, after the calibration, all capacities of the lockers are rounded up to a multiple of 20. This results in 3012 self-collection points in the study area with a total capacity of 60240.

5.2.2. Sensitivity Analysis

In this section, a sensitivity analysis will be performed for different modelling factors, which is done to understand the effects of the model choices better and to validate those choices. The used factors are selected because of their plausible impact on the results and because most are based on research output or assumptions instead of revealed outcomes. First, the beta weights of the attributes used in the performance score will be analysed to provide insight into when different delivery methods are more likely to be chosen and to interpret the elasticity of the attributes. Secondly, different magnitudes for WoM and familiarity will be evaluated. Then, the impact of innovation on the preference probability of drones is assessed. Lastly, different capacity evolution rules are compared to test their impact.

Attribute Beta Weights

The performance score of a delivery method depends on the beta weights of each attribute, whereby the magnitude of the beta weights and the relative weight compared to the other beta weights are important. Since empirical research on consumer preferences and choice for last-mile delivery attributes is quite limited, and the results are not always consistent general estimations for the beta weights of the attributes do not exist. Seven cases of beta weights are tested in scenario three, thus with van, self-collection, crowdshipping and drone delivery. However, the innovation of drone possibility is not applied to make the runs of all cases comparable. For each case, three runs are realised to check for consistency. The different cases can be seen in Table 5.3 and are elaborated below.

Case 1: Equal: All attributes are set to the same weight of 1, representing a neutral sensitivity towards all attributes.

Case 2: Speed insensitive: Although consumers prefer a fast delivery, multiple studies indicate that they are willing to wait longer if the delivery is cheaper, more convenient or, for example, smaller delivery windows are offered (Caspersen & Navrud, 2021) (Nguyen et al., 2019) (Buldeo Rai et al., 2019). Therefore it can be assumed that consumers are insensitive to the delivery speed of a parcel.

Case 3 & 4: Costs sensitive: In general, delivery costs are seen as the most influential attributes for a consumer's utility. In Nguyen et al. (2019), the cost attribute is six times more important than all other

Table 5.3: Cases for attribute beta weights

Attribute	**Case 1** Neutral	**Case 2** Speed Insensitive	**Case 3** Costs Sensitive	**Case 4** Highly Costs Sensitive	**Case 5** Reliability Sensitive	**Case 6** Highly Reliability Sensitive	**Case 7** Distance Insensitive
β_{Speed}	1	0.5	1	1	1	1	1
β_{Costs}	1	1	2	4	1	1	1
$\beta_{Reliability}$	1	1	1	1	2	4	1
$\beta_{pick-upDistance}$	1	1	1	1	1	1	0.5

attributes like delivery speed and a time slot. This is confirmed by Buldeo Rai et al., who concluded that delivery price is almost four times as important to consumers as delivery speed. The results of Maltese et al. (2021) also showed that costs are twice as important to consumers than the delivery speed in the case of grocery shopping. Thus, the costs attribute should likely be the attribute with the highest beta weight. In cases 2 and 3, the extent of the cost sensitivity is examined.

Case 5 & 6: Reliability sensitive: To the writer's knowledge, no stated preference studies use reliability as an attribute with the possibility of a failed delivery. The disutility of delayed delivery is very moderate (Caspersen & Navrud, 2021), which aligns with the insensitivity for delivery speed. However, it can be expected that consumers will perceive a very high inconvenience when a parcel is not delivered. Therefore, in these cases, the reliability beta is given a weight of 2 and 4.

Case 7: pick-up distance sensitive: Molin et al. (2022) presents that the utility for increasing pick-up distance is negative, nevertheless considerably weaker than the effect of rising delivery costs. This indicates that consumers could be relatively insensitive to the distance they need to travel to retrieve a parcel.

Results Beta Sensitivity

In Figure 5.5, the average performance score of all delivery methods across all the iterations of a beta weight case can be seen. Based on those results the chosen beta weights are 1, 2, 2 and 1. It can be concluded that the performance score behaves as expected. An increase in a beta value corresponds with a higher performance score (thus a worse operational performance); the opposite is true for reducing a beta weight. The model is most sensitive for the reliability attribute, as doubling the beta weight increases the average performance score for all delivery methods by 42%. The cost attribute follows with a change of 25%, speed with -11% and distance with -5%. Additionally, other beta weights for reliability and costs widen the difference between the minimum value the delivery methods reach and the maximum value.

The probability of preferring van delivery is slightly changed in cases 2, 5 and 7 compared to case 1; see Figure 5.6. An insensitiveness towards delivery speed makes van delivery more popular, as it is one of the slower delivery methods. An increase in the sensitivity for reliability makes the probability of preferring van less likely, a consequence of the failed delivery rate of 25%. Furthermore, van delivery is not affected in its performance score by the distance score. However, the lower weight of the distance attribute reduces the performance score of self-collection, consequently lowering the probability of preferring van delivery as self-collection becomes more competitive.

An increase in the costs beta (cases 3 and 4) significantly affects the probability of preferring van delivery. Because van delivery is free, preferring this delivery method becomes much more likely due to the performance scores of crowdshipping and drone delivery worsening to a higher value.

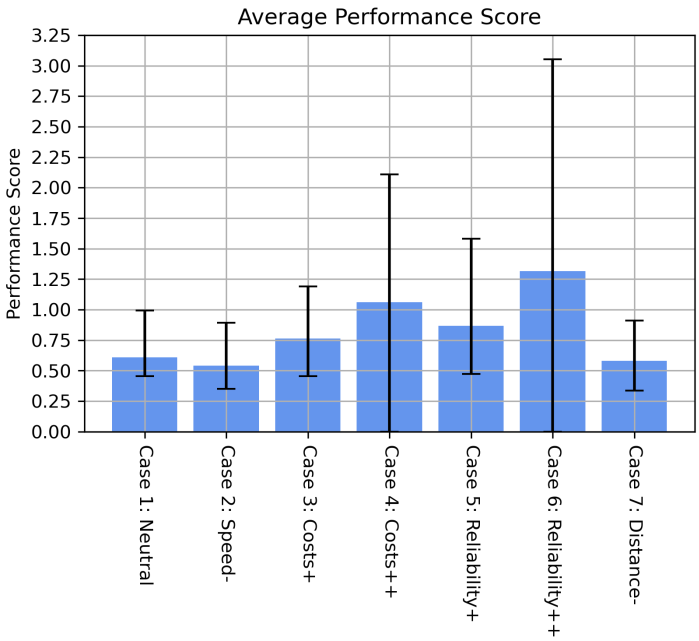
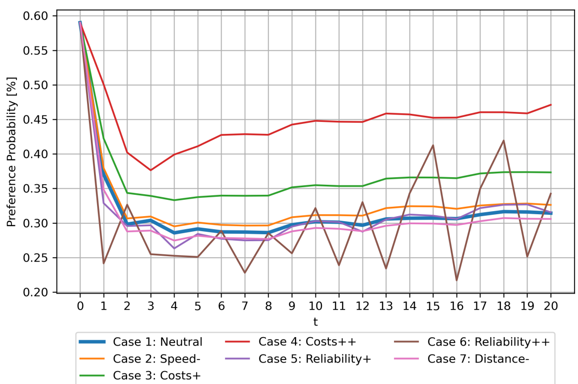
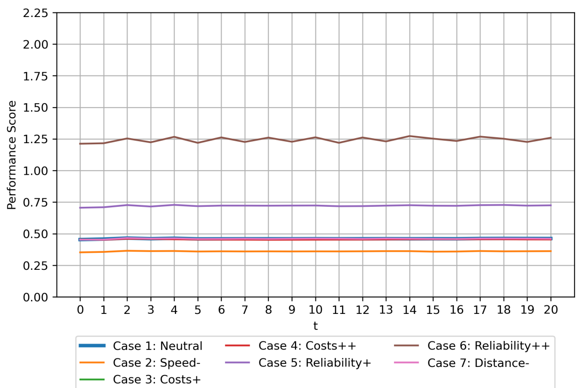


Figure 5.5: The average performance score over all iterations of all delivery methods in each beta case



(a) Probability Van with different beta weights

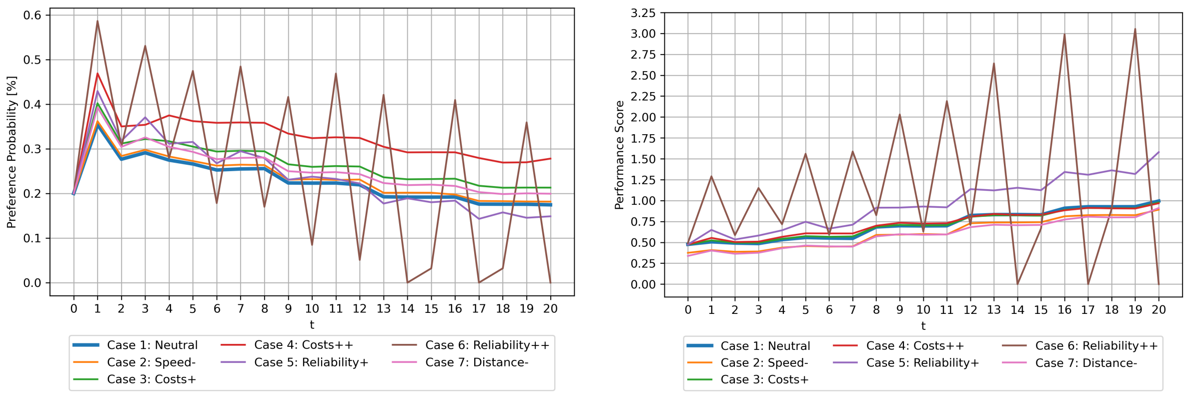


(b) Performance Score Van with different beta weights

Figure 5.6: Performance van delivery during the beta sensitivity analysis

A beta weight for reliability of four (case 6) does affect van delivery in two ways. First of all, the performance score increases substantially, resulting in a lower probability of preferring van delivery. Secondly, the probability of preferring van delivery becomes very unstable, while the performance score is rather constant. This is caused by the highly alternating performance scores of self-collection and drone delivery; see Figure 5.7 and Figure 5.9. With self-collection, the following dynamic occurs: at the first time point, the likelihood of preferring self-collection is high, considering the other delivery methods perform worse. In the next iteration, the demand for parcels exceeds the supply, resulting in a low reliability and, thus, a high reliability score. The peaks become larger over time since the capacity of self-collection points reduces each year due to the shrinking demand. In the case of drone delivery, a similar dynamic takes place; however, the effect is lessened by the increase of drones each half year. As these delivery methods fluctuate simultaneously, the probability of preferring van and crowdshipping develops changeability.

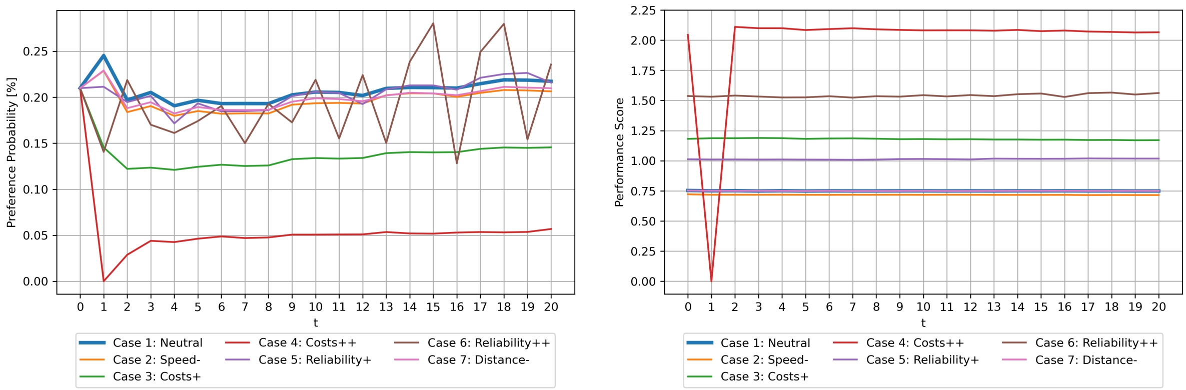
The beta for distance does only adjust the performance score of self-collection. The impact of this change is limited as at $t = 20$, the probability of preferring self-collection is around 17.5% in case 1 and just under 20% in case 2. Additionally, the performance score does not improve significantly.



(a) Probability Self-Collection with different beta weights

(b) Performance Score Self-Collection with different beta weights

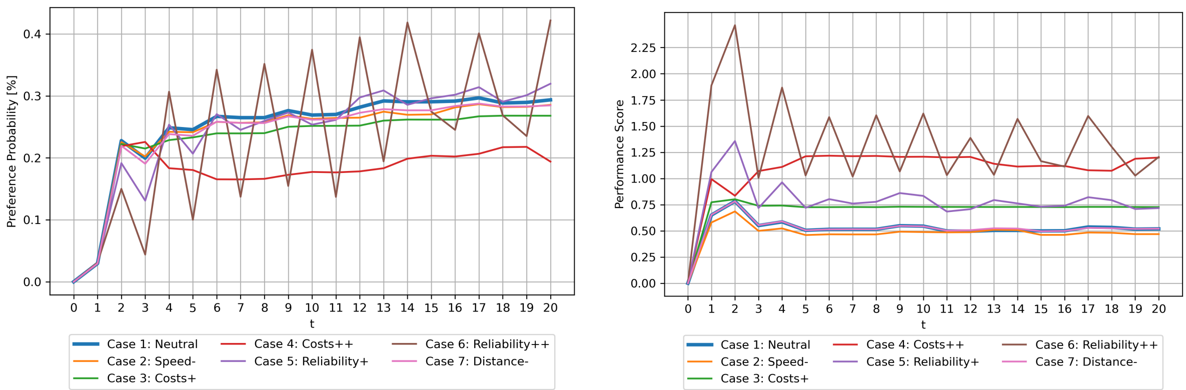
Figure 5.7: Performance self-collection delivery during the beta sensitivity analysis



(a) Probability Crowdshipping with different beta weights

(b) Performance Score Crowdshipping with different beta weights

Figure 5.8: Performance crowdshipping delivery during the beta sensitivity analysis



(a) Probability Drone with different beta weights

(b) Performance Score Drone with different beta weights

Figure 5.9: Performance drone delivery during the beta sensitivity analysis

Crowdshipping is very dependent on the beta weight for costs, being the most expensive delivery method. With a beta weight of four, the probability of preferring crowdshipping even becomes zero at $t = 1$. After that, the probability turns stable, however, at a much lower level. At that time point $t = 1$, the performance score drops to zero due to delivering zero parcels with crowdshippers.

To conclude, the developed model is, to a certain extent, insensitive to the speed and distance attribute. On the other hand, the beta weights for costs and reliability significantly impact the probability and operation of the four delivery methods. Considering these findings, the beta weight for speed and pick-up distance will be set to 1. As there are strong indications in the literature that consumers are costs sensitive and the model's predictions are impacted heavily by this attribute, the weight for costs is assumed to be 2. Literature generally presents that delivery costs are twice as important as other attributes; on that ground, a weight of 4 is too extensive. Although earlier research did not explore a reliability attribute, certainly with the possibility of a failed delivery, it is assumed that consumers are sensitive to a delivery that does not succeed. Thereby this model indicates that the operation and preference of a delivery method have a strong interaction. Thus the beta weight of reliability is assumed to be 2.

Table 5.4: The attribute weights used in the simulation

Attribute	Weight
β_{Speed}	1
β_{Costs}	2
$\beta_{Reliability}$	2
$\beta_{pick-upDistance}$	1

Word of Mouth

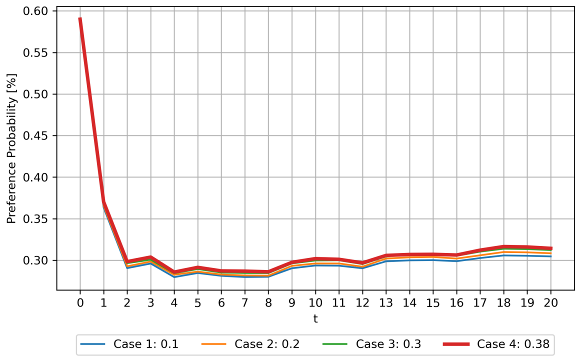
Mahajan et al. (1995) propose that the coefficient of imitation in Equation 4.9, q, is generally 0.38 per year and ranges between 0.3 and 0.5. In this simulation model, the WoM effects take place each quarter. Therefore it is assumed that a value of 0.38 for q is an overestimation of the magnitude of the WoM effect (Massiani & Gohs, 2015). However, typical quarterly values for q in the context of a parcel delivery service are unknown. Therefore, a sensitivity analysis for different values of q is performed. The tested values can be seen in Table 5.5. The sensitivity analysis will be performed similarly to that of the beta weights. Thus scenario three without the implementation of drone innovation. All betas are set to one.

Table 5.5: Cases for the imitation coefficient

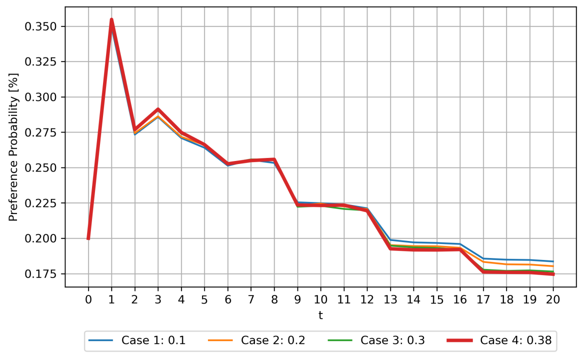
	Case 1	Case 2	Case 3	Case 4
q	0.1	0.2	0.3	0.38

The system sensitivity for the imitators coefficient is relatively low. The preference probability of the delivery methods changes only slightly, as seen in Figure 5.10. The largest relative variation appears at $t = 6$ with crowdshipping, where there is a difference of 6% between case 1 and case 4. Also, the highest standard deviation can be found with crowdshipping; at $t = 2$, this measure reaches 0.0048.

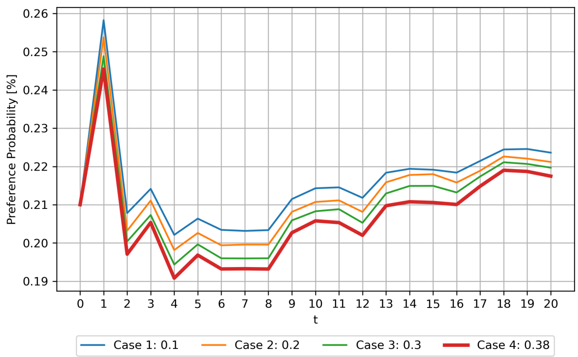
The modest sensitivity for change in q results from the low magnitude of the WoM effect, see Figure 5.11. In case 4, where WoM has the most substantial result, the magnitude ranges between -0.026 and 0.016, as the magnitude of drone delivery at $t = 0$ arises from the innovation coefficient for new delivery methods. An important reason for the low magnitude is that the magnitude scales with the performance scores of the delivery methods. Consequently, the WoM magnitude projected by the Bass diffusion model is multiplied by, for example, an absolute value between 0 and 0.32 in case 1. Because of this downsizing effect, the coefficient of imitation will be maintained at 0.38.



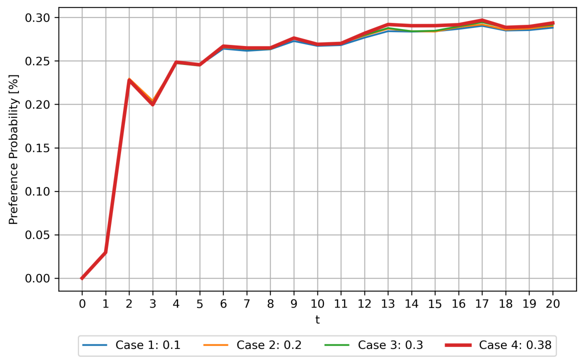
(a) Probability Van with different q



(b) Probability Self-Collection with different q



(c) Probability Crowdshipping with different q



(d) Probability Drone with different q

Figure 5.10: Probability of the delivery methods during the WoM sensitivity analysis

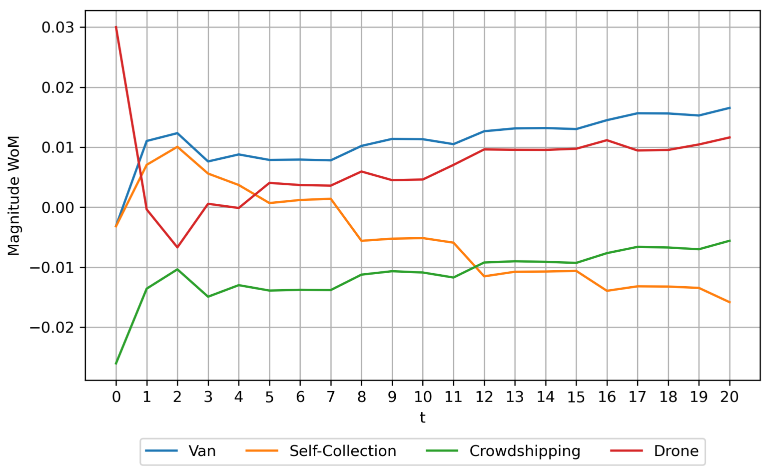


Figure 5.11: Magnitude of WoM in case 4

Familiarity

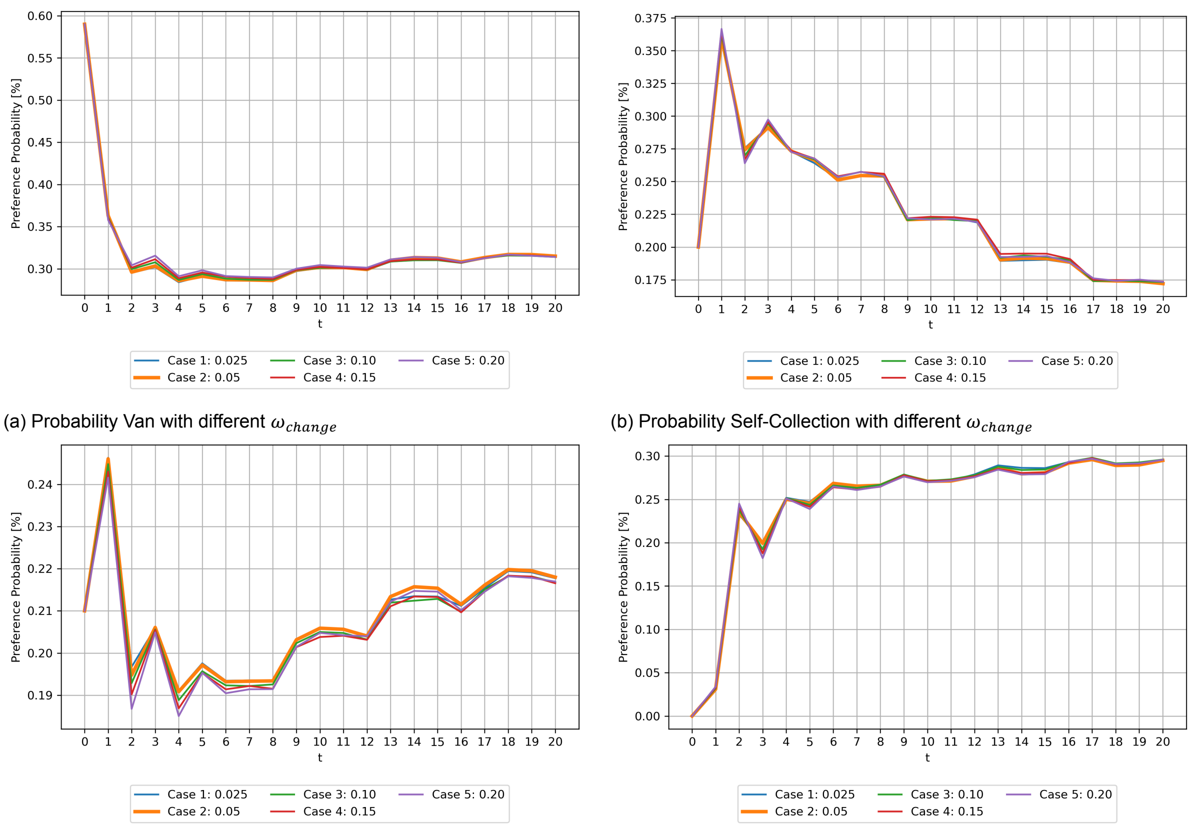
As indicated in subsubsection 4.5.4.6, a new method to estimate the familiarity effect is developed for this study. Therefore the magnitude and the sensitivity of this effect are evaluated. In Table 5.6, the used values for ω_{change} can be seen. The sensitivity analysis is executed similarly to the WoM sensitivity analysis.

Table 5.6: Cases for the resistance value

	Case 1	Case 2	Case 3	Case 4	Case 5
ω_{change}	0.025	0.05	0.1	0.15	0.2

Familiarity impacts, on a small scale, the preference probability of all the delivery methods, as seen in Figure 5.12. Van delivery starts with the highest probability; therefore, the familiarity effect is lowest on that mode, see Figure 5.13. Consequently, with higher values of ω_{change}, the probability of preferring van delivery is slightly higher. This effect is most substantial in the first eight runs. The largest relative difference appears with drone at $t = 2$, where the probability of preferring drone is 9% higher in case 1 than in case 2. With van delivery, the highest standard deviation occurs at $t = 1$ with a value of 0.0067. The differences between self-collection and crowdshipping stay between -2% and 3%. The familiarity effect shrinks to zero in the last iterations, and all the probabilities converge to a similar outcome.

It is assumed that familiarity's impact is comparable with WoM's. Therefore, ω_{change} is set to 0.15 because it results in a similar magnitude.



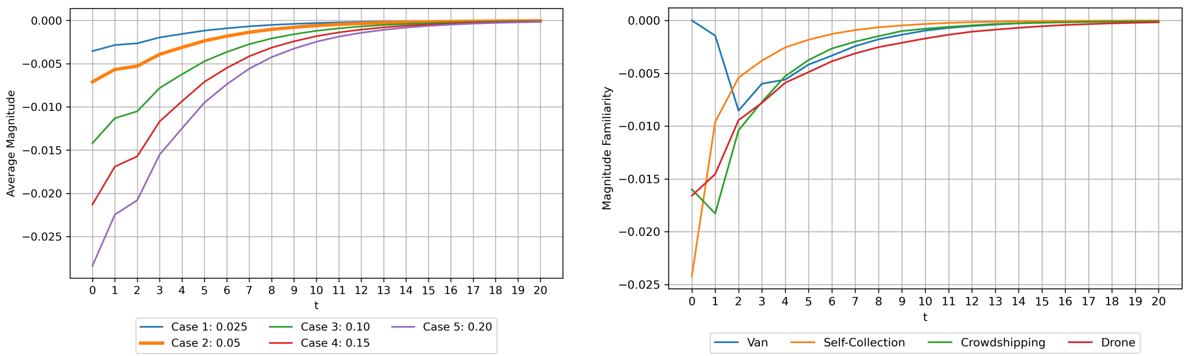
(a) Probability Van with different ω_{change}

(b) Probability Self-Collection with different ω_{change}

(c) Probability Crowdshipping with different ω_{change}

(d) Probability Drone with different ω_{change}

Figure 5.12: Probability of the delivery methods during the familiarity sensitivity analysis



(a) The average magnitude of familiarity during the familiarity sensitivity analysis

(b) Magnitude of familiarity in case 3

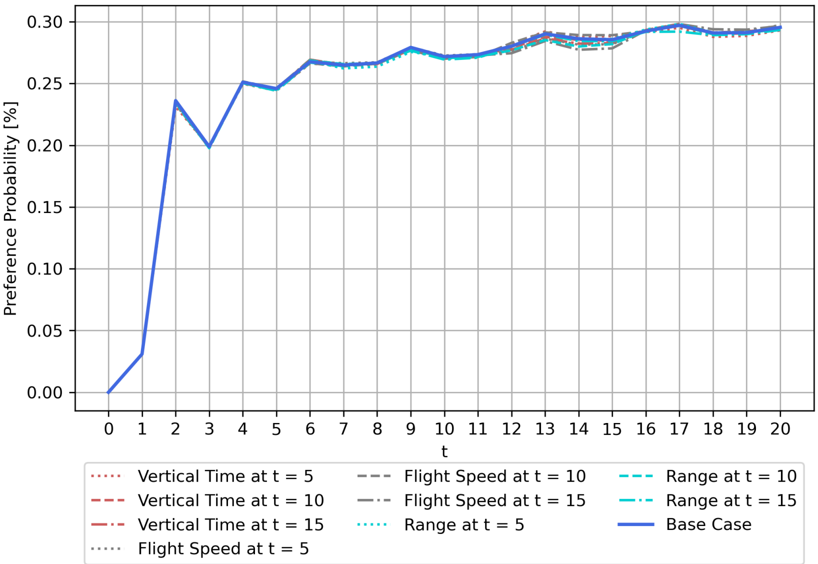
Figure 5.13: Magnitude of familiarity

Innovation

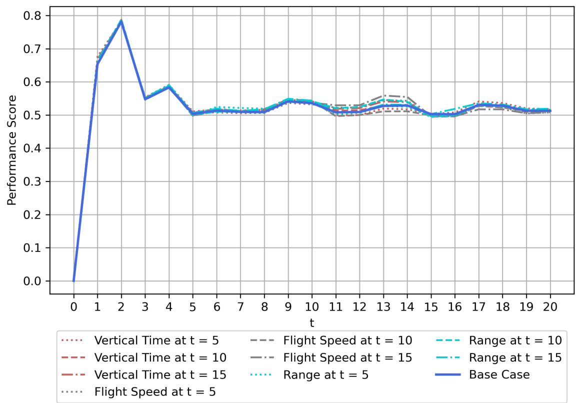
This sensitivity analysis looks into the effect a drone innovation can have. There are three possible innovations: 1) the vertical travel time is reduced; 2) the average flight speed is improved; 3) the range is increased. For each innovation, a simulation is performed where an innovation takes place at $t = 5, 10\ or\ 15$.

The results on the preference probability of drone are shown in Figure 5.14. It can be concluded that the innovations do not proceed significantly with other probabilities for drone delivery. This is also reflected by the performance scores that do not improve due to the innovations. The differences resulting from the reduced vertical flight time are the smallest, indicating that this innovation is not a limiting factor in drone operations. It was theorized that increased flight speed would allow the drones to deliver more parcels within the time frame. This is, however, not reflected in the reliability score that increases in the case of $t = 5$ and $t = 15$, while it decreases with $t = 10$. A similar effect can be seen in the speed score. The innovation of increased flight range does have a sensible impact. By increasing the range, the reliability and speed score do decrease slightly, representing an improvement. As the delivery cost depends on the flight distance, the costs score increases. As the beta weights are equal at one, it could be that with other weights, the preferences for drone would shift because of this innovation.

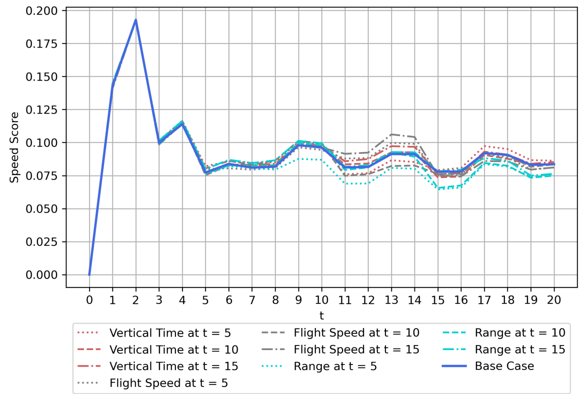
Because the influence of vertical time and flight speed innovations is very limited, those innovations will not be implemented during the final simulations. The possibility of a range extension will be implemented.



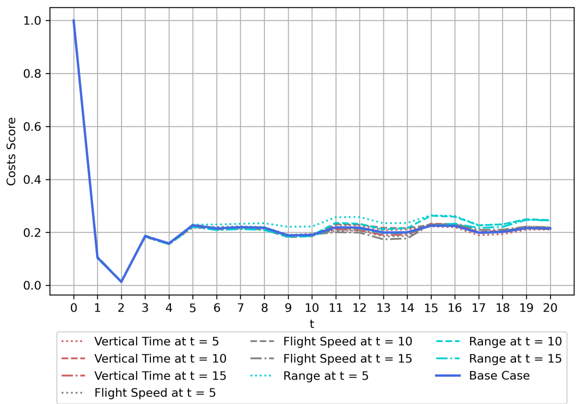
(a) The probability of drone with different innovations at various time points



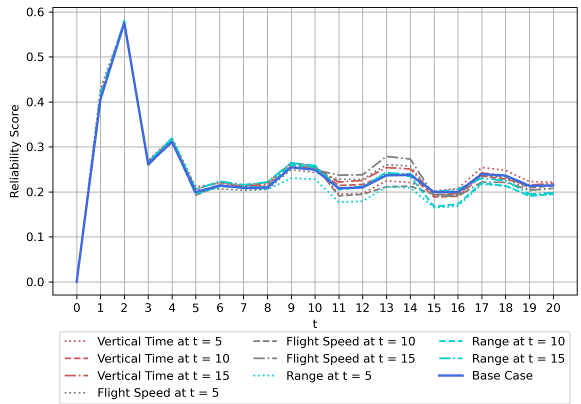
(b) Total performance score drone with different innovations at various time points



(c) Speed score drone with different innovations at various time points



(d) Costs score drone with different innovations at various time points



(e) Reliability score drone with different innovations at various time points

Figure 5.14: Results of the innovation sensitivity analysis

Capacity Growth
The number of self-collection points per zone and the number of drones per carrier depot threshold values determine if points or drones are added or removed. In this test, the sensitivity of those thresholds is analysed. Six cases are evaluated in the case of self-collection points, see Table 5.7, with case 3 being the base case. For the drone threshold, four cases are assessed; see Table 5.8.

Table 5.7: Cases for the threshold values for self-collection evolution

	Case 1	**Case 2**	**Case 3**	**Case 4**	**Case 5**	**Case 6**
Add self-collection	30%	40%	50%	60%	50%	50%
Remove self-collection	75%	75%	75%	75%	50%	60%

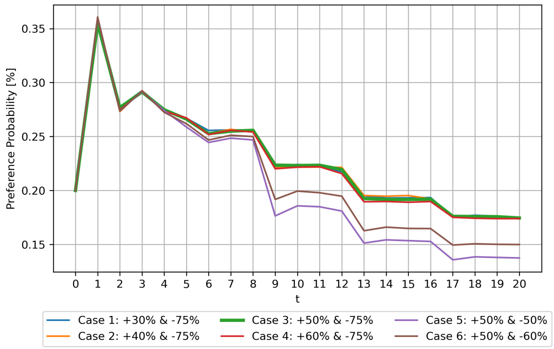
Table 5.8: Cases for the threshold values for drone evolution

	Case 1	**Case 2**	**Case 3**	**Case 4**
Add a drone	15%	25%	35%	45%
Remove a drone	15%	25%	35%	45%

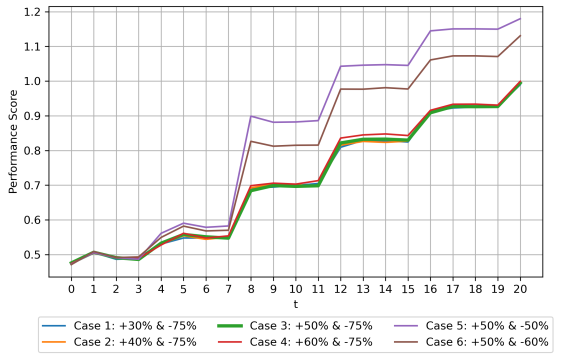
From the results shown in Figure 5.15, it can be concluded that self-collection evolution is insensitive to the threshold of adding a point and sensitive to the removal threshold. The insensitivity to the growing threshold results from the low self-collection demand; thus, the removal effect is dominant. In a scenario where the demand is higher, it could be that the sensitivity is comparable with the sensitivity of the removal threshold. That threshold greatly impacts the number of self-collection points, consequently deteriorating the reliability and distance to a higher score and, thus, the total performance score. Over time the preference probability for self-collection drops significantly, reducing the demand even further, regardless of the growth of the total parcel demand.

It is assumed that a carrier will not remove a self-collection point very quickly because the investment costs are high, while operational costs are relatively low. Therefore, it can be hypothesised that the removal threshold can be set to a high value. Following that argumentation, it is expected that the self-collection points will not be added unless there is a clear indication of their added value. Accordingly, the threshold for adding a self-collection point is set at 50% and for removing at 75%.

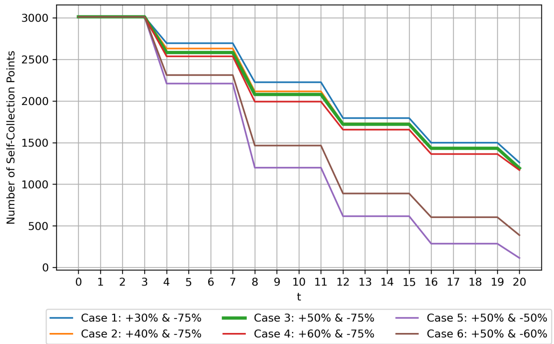
The results of implementing different evolution thresholds for drones are shown in Figure 5.16. The results of all cases are highly comparable and have a similar outcome at the time horizon. Both the preference probability and performance score of drone converge after six iterations, even with a constantly growing demand due to the yearly growth in parcel demand. The only noteworthy deviation can be seen in the number of drones at the depots between $t = 10$ and $t = 13$. Here a slight difference in the demand/capacity ratio leads to a significant variation in the number of drones. As both cases 1 and 4 have a smaller growth, it cannot be confirmed that different thresholds are more responsive. At $t = 13$, the vast parcel demand growth equals the number of drones again. Concludingly, the model is insensitive to the drone evolution threshold, mainly to the enormous growth in demand. Thus the threshold is kept at 25%.



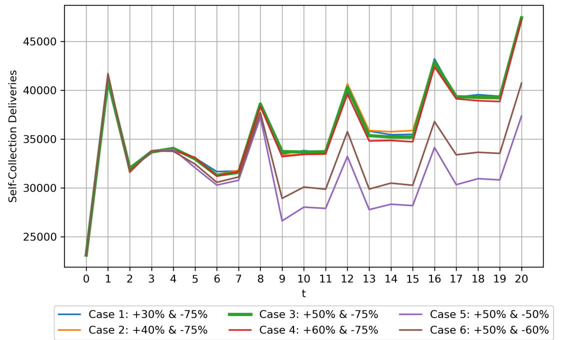
(a) Probability Van with different evolution thresholds



(b) Total performance score Self-Collection with different evolution thresholds

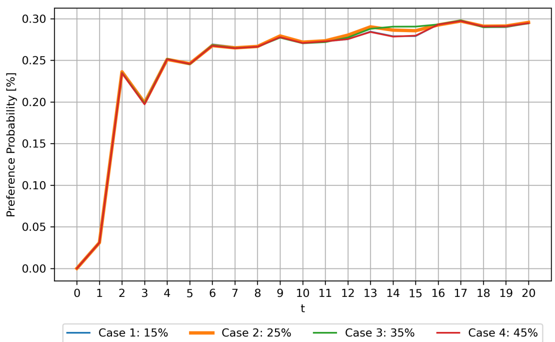


(c) Total number of self-collection points with different evolution thresholds

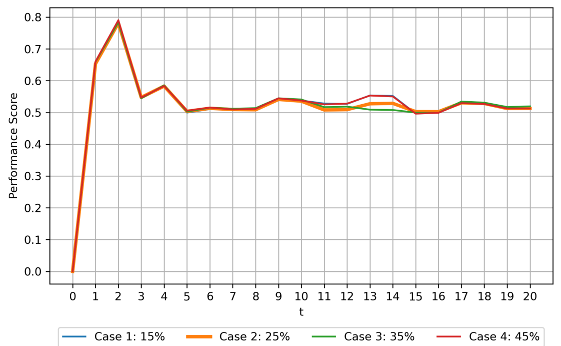


(d) Total number of self-collection deliveries per day with different evolution thresholds

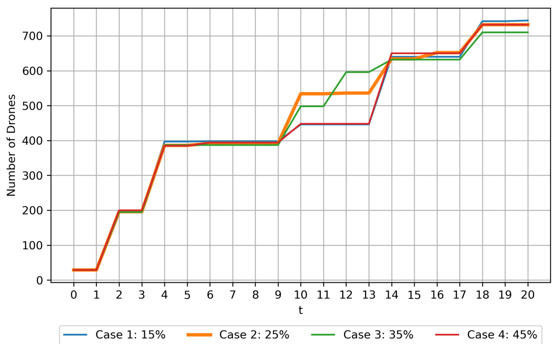
Figure 5.15: Results of the growth sensitivity analysis for the self-collection evolution threshold



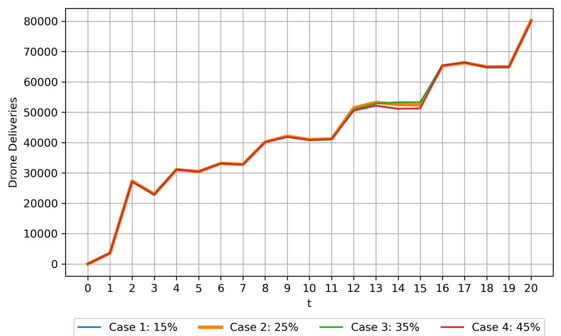
(a) Probability Drone with different evolution thresholds



(b) Total performance score Self-Collection with different evolution thresholds



(c) Total number of drones at the depots with different evolution thresholds



(d) Total number of drone deliveries per day with different evolution thresholds

Figure 5.16: Results of the growth sensitivity analysis for the drone evolution threshold

5.2.3. Verification & Validation

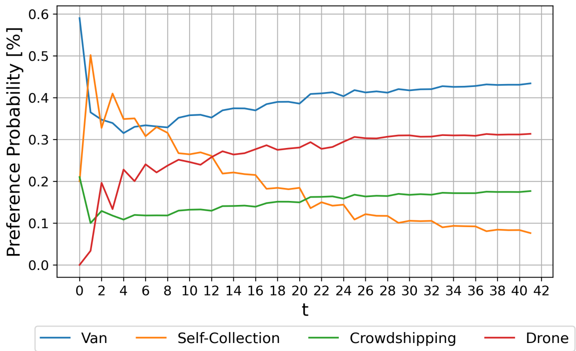
Verification is carried out to ensure that the developed model is correct and matches the model specifications (Carson, 2002). By validating the model, it can be established that it represents the actual system sufficiently. To verify and validate the proposed approach, three tests are performed: 1) Test for face validity; 2) Test for various input parameters; 3) Compare model predictions with the performance of the actual system or with predictions from other studies (Sargent, 2010).

Verification

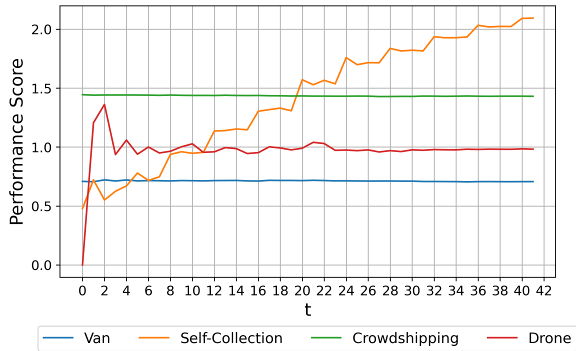
The sensitivity analysis already provided extensive insight into the model behaviour. In all cases, irrespective of the sensitivity, the model output changes in a logical manner that corresponds with the theory. This indicates that the developed model functions appropriately and can be used to evaluate the interactions within the system.

In addition, the sensitivity analysis tested the model for a range of input parameters. The results showed that the direction and magnitude of output changes. Only in the case of an extreme value for the beta of reliability, the model showed unstable behaviour; however, even in that case, the results were explainable.

To further verify the model, a time horizon test is performed. The time horizon test is carried out to verify that the model behaviour is consistent outside the time frame for which the model is designed. The results from the time horizon test can be seen in Figure 5.17. The behaviour of the model becomes relatively constant with a longer simulation horizon. Except for self-collection, the preference probabilities and performance scores are stable in the second half of the simulation. Therefore, it can be concluded that the simulation model also performs for longer time periods and that there are no sudden changes after the cut-off point of this study.



(a) Preference probabilities in a time horizon test



(b) Total performance scores in a time horizon test

Figure 5.17: Results time horizon test

Validation

To validate the model, the output can be compared with the real world and other research. In Table 5.9, the market shares estimations from the model and presented in literature are compared. The values for van delivery are the currently operated percentage in the Netherlands. The market shares for self-collection, crowdshipping and drone are estimation results based on stated preference studies.

Scenario 1 is comparable with the current last-mile environment in the Netherlands. In 2021 and 2022, around 80% of the parcels were delivered to a consumer's home in the Netherlands. For scenario 1, that means roughly 80% of the parcel should be delivered by vans to reflect the current situation. The model estimate is significantly lower and closer to the estimates of Molin et al. (2022). That study estimates that with the current self-collection infrastructure, up to 29% of the consumers would prefer self-collection. With a denser network, that percentage could increase considerably. Thus the model estimates of scenario 1 overestimate the self-collection market share of the real world. However, the shares of ACM reflect the performed delivery and not the preferred delivery. Moreover, the results align with predictions in the literature.

Table 5.9: Market share model estimation versus literature

Market share	Value at t=20			Literature value	Scource
	Scenario 1	Scenario 2	Scenario 3		
Van	69%	53%	39%	79.5 - 81.5%	ACM, 2021 & ACM, 2022
Self-collection	31%	25%	18%	18 - 29%	Buijs and Niemeijer, 2022 & Molin et al., 2022
Crowdshipping		22%	15%	21 - 27%	Buldeo Rai et al., 2021
Drone			28%	7.18 - 53%	Kim, 2020 & Merkert et al., 2022

The model estimate for consumer preferences of crowdshipping does also align with the research of Buldeo Rai et al. (2021), although that study did estimate the share of crowdshipping in a case with only van delivery. Both Merkert et al. (2022) and Kim (2020) describe that the potential market share of drone delivery can deviate largely. For low-value parcels and high delivery fees, only a limited share of consumers will choose drone delivery, while with important parcels, consumers are more likely to prefer drones. The model estimation lies in the middle of this range, which the moderate delivery costs could explain. Thereby, this model does not account for parcel value.

Because the model results are, to a large extent, in line with other research, it can be concluded that the developed model can be used to answer the research questions. Even within the literature, there is strong variation in market share predictions. Therefore very accurate validation is complicated. However, the model behaviour does reflect the real-world system and can thus be used to analyse the last-mile system.

5.3. Scenario Analysis

In this section, the results from the simulation model are presented. The results from the three scenarios will first be discussed separately and will then be compared. In all scenarios, one run will consist of 21 iterations called t, and each iteration represents a time step of a quarter of a year. Thus, the entire simulation has a time horizon of 5 years.

5.3.1. Scenario 1: Van & Self-Collection

This scenario reflects the current last-mile operation in the Netherlands with the possibility of van delivery at the consumers' home or delivery to a self-collection point. In Figure 5.18, the average preference probability of both delivery methods is shown at each simulation step. Five runs of each 21 iterations are performed. The bars reflect the minimum and maximum value of the five runs at that iteration. The stochasticity in the model produces little difference between various runs.

In the first four iterations, the demand for both delivery methods develops to a share of 50%. From $t = 3$, van delivery becomes the generally preferred delivery method, growing to around 70% at the time horizon. The shrinking preference probability for self-collection results from the increasing performance score, which represents a deteriorating service quality, see Figure 5.19. While the performance score of van delivery is constant at around 0.7, the performance score of self-collection increases due to the worsening reliability score. The reliability score is increasing because of two effects: 1) Although the probability for self-collection declines, the absolute demand advances from roughly 55.000 parcels per day to 75.000, see Figure 5.20a. This puts a larger burden on the available capacity. 2) On the contrary, the total supplied capacity declines over time; see Figure 5.21. The decline follows from the threshold that dictates when a self-collection must be kept, removed or added in a zone.

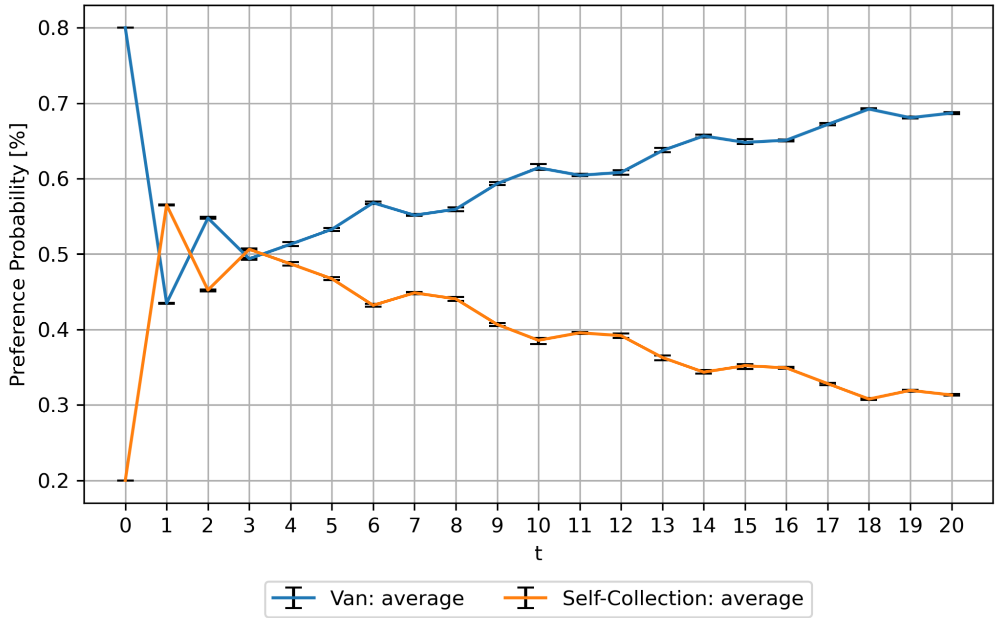
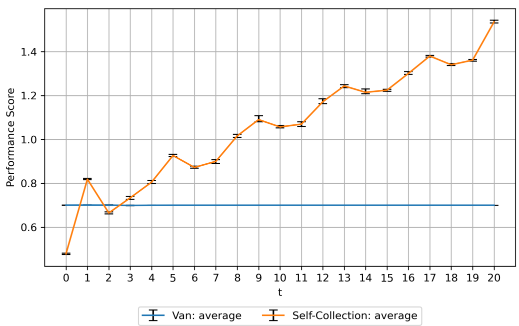
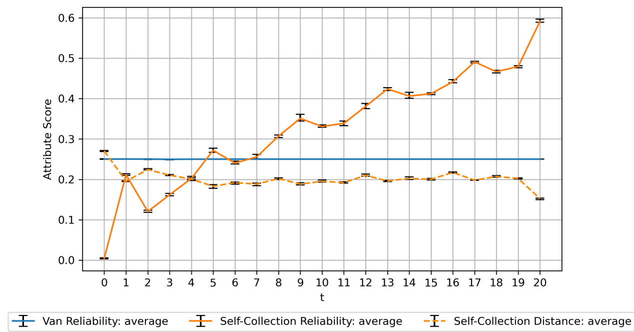


Figure 5.18: Method preference probability in scenario 1

The growth in van demand is reflected by the rise in the number of daily trips, as shown in Figure 5.20b. While the preference probability for van rises from 50% at the start to around 69% at the time horizon, an increase of around 38%, the delivery demand grows from roughly 92.500 to 155.000, an increase of 67.5%. This is due to the yearly growth of the parcel demand. To fulfil the increasing demand also, the number of trips that need to be performed inflates likewise; see Figure 5.20b. An important notion is that 25% of the parcel trips could have a failed delivery. And thus, in real-world operation, those parcels need to be reassigned to a delivery method, consequently increasing the number of required deliveries even further. For self-collection delivery, this is not the case because more parcels are dropped at each self-collection point over time. In the iteration at $t = 0$, only 2 parcels are dropped per stop at a self-collection point on average, while at $t = 20$, this has grown to more than 8.5 parcels on average per stop at a self-collection point.

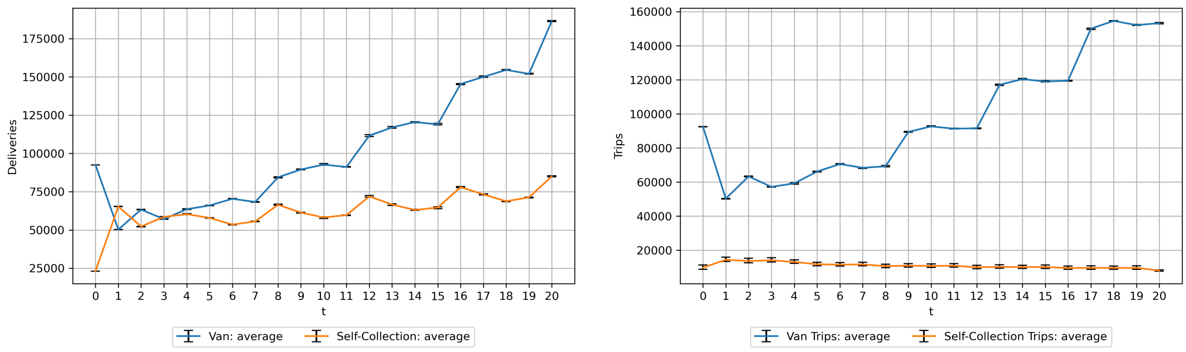


(a) Performance scores



(b) Average attribute scores for distinctive attributes

Figure 5.19: Performance van and self-collection in scenario 1



(a) Daily number of parcel requests

(b) Daily number of trips

Figure 5.20: Daily parcel requests and trips in scenario 1

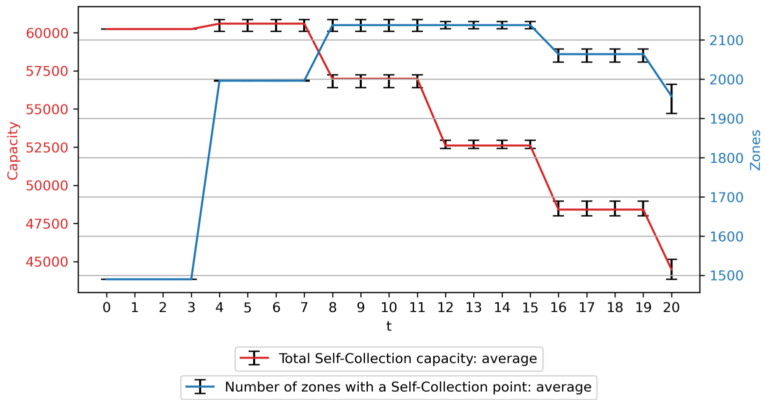
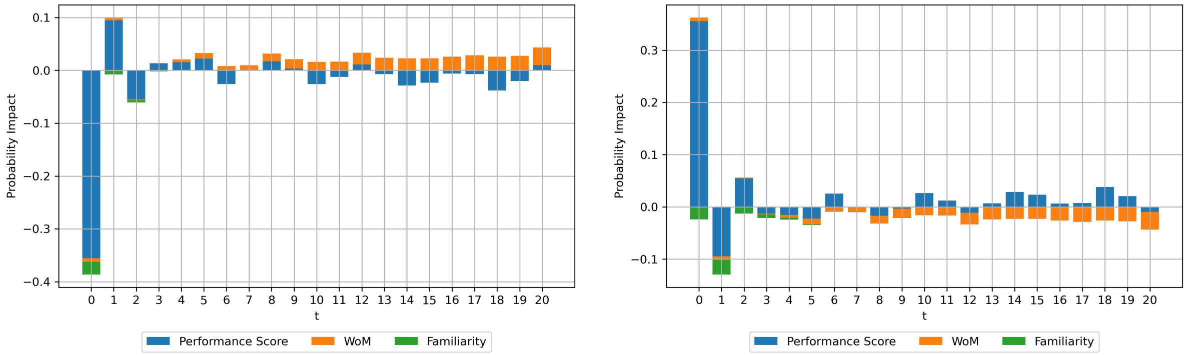


Figure 5.21: Evolution of capacity and distribution self-collection points in scenario 1

Table 5.10: Vehicle kilometres and emissions vans for zone to zone and intrazonal travel in scenario 1

Travel	**t=0**				**t=20**			
	VKMs	VKMs %	kg CO2	CO2 %	VKMs	VKMs %	kg CO2	CO2 %
Zone-to-Zone	66213	73%	12511	68%	69753	64%	13261	58%
Intrazonal	24181	27%	5972	32%	39040	36%	9643	42%
Total	90394		18484		108793		22904	

To fulfil all requests, vans have to drive to each individual location of a van request and to all self-collection points with a parcel demand. In Table 5.10, the corresponding vehicle kilometres and emissions are shown. A distinction is made between travel between zones and within a zone. From that, it can be concluded that travel from zone to zone is the main contribution to the distance and emissions. However, the intrazonal emissions are relatively higher. This is because it is assumed that intrazonal travel takes place on city roads, where the emission factor (CO2 g/km) is high. The total vehicle kilometres grows by just 20%, compared to a total parcel demand growth of 93%. Additionally, zone-to-zone travel increases by only 5.3%, while intrazonal travel increases by 61%. The zone-to-zone vehicle kilometres change only slightly because, although there are fewer locations to stop, for example, iteration 0 compared with 20, still most of the zones need to be visited. With growing demand, the number of zones with demand stays quite constant. However, the number of stops within each zone increases significantly. Consequently, at $t = 0$, the average distance driven to deliver a parcel is 0.78 km, while this shrinks to 0.49 km at $t = 20$. Correspondingly, the CO2 emission per parcel dropped by 36% because of the higher efficiency. Nevertheless, due to the increasing parcel demand, the total CO2 emissions grow by roughly 24% in the simulation horizon of five years. Thereby, the share of vehicle kilometres and emissions that are made within zones, and thus often at low-capacity roads in urban areas, increase heavily.



(a) Impact on the preference probability for van

(b) Impact on the preference probability for self-collection

Figure 5.22: Preference probability impact in run 1 in scenario 1

Three factors are modelled that influence consumer preferences. In Figure 5.22, the impact on the preference probability per factor is shown for van and self-collection in the example of run 1. The first factor is the performance score, which is of significant influence in the first iterations. Since the performance scores of both delivery methods are pretty similar at the beginning, the logit formulation estimates that the preference probabilities will correspondingly be quite similar. Consequently, the performance factor pulls the preferences closer to each other. When the probability reflects the logit estimation, the impact reduces.

The WoM impact increases over time due to the growing discrepancy between the performance scores. It can be seen that WoM becomes the dominant factor in later iterations and thus is the reason for a rising probability of preference probability, despite the logit formulation estimating a lower probability. Lastly, the familiarity factor always has a declining effect. The magnitude of that effect quickly converges to zero. As self-collection delivery is the lesser-known delivery method, the familiarity effect is relatively more substantial for that delivery method and absolutely larger from timepoint $t = 1$.

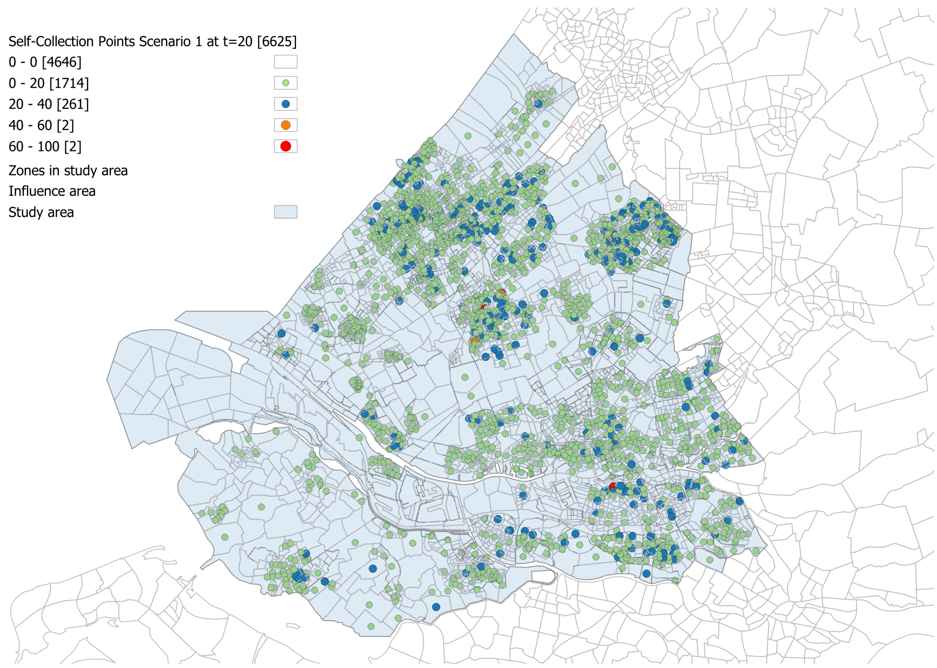
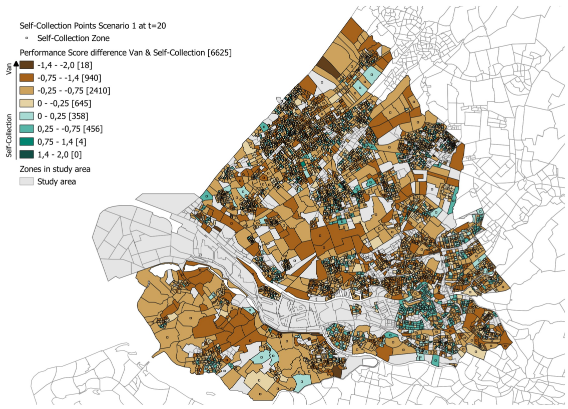
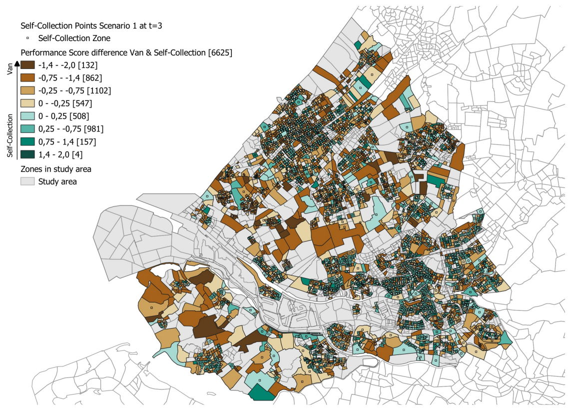


Figure 5.23: Self-Collection point distribution across the study area at $t = 20$ in scenario 1

As described earlier in this subsection, the capacity of the self-collection points reduces during the simulation. On the other hand, the number of zones with a self-collection point increases. This can also be seen in Figure 5.23, where only four zones possess three or four points and 261 with two points. At the start, there were 1474 zones with two self-collection points and 16 with three points; see Figure 5.3. Compared with the initial distribution, self-collection points are, to some extent, moved from highly urbanized areas, like The Hague and Rotterdam, to smaller cities and villages. This redistribution occurs because the model seeks to reduce the distance score, which should result in higher service quality. Because of this, the model prioritises adding a self-collection point to a zone without such a point above having one zone with a high self-collection capacity. At $t = 20$ in, for example, the city of the Hague, almost every zone possesses a self-collection point. Consequently, self-collection points that are removed from a zone with multiple points are placed in less urbanized areas, like a village. Currently, in the Netherlands, parcel carriers are expanding the supply of self-collection points, mainly automated lockers. Thus the model behaviour is in line with that development. On the other hand, many self-collection points are removed by the model, resulting in a worsening reliability score. Currently, the Dutch carriers are generally not moving self-collection points. Instead, they are adding them (ACM, 2020) (ACM, 2021). Thus the model result is not exactly in line with reality.



(a) Average performance score difference van and self-collection

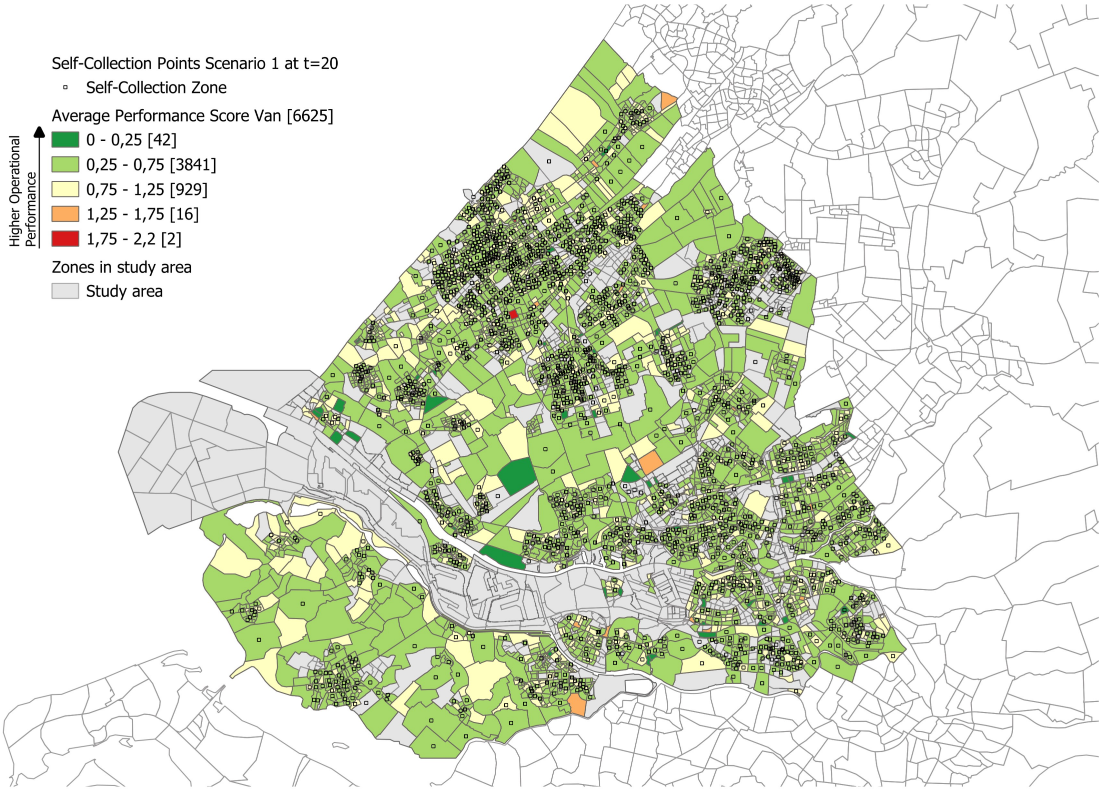


(b) Performance score difference van and self-Collection at $t = 3$

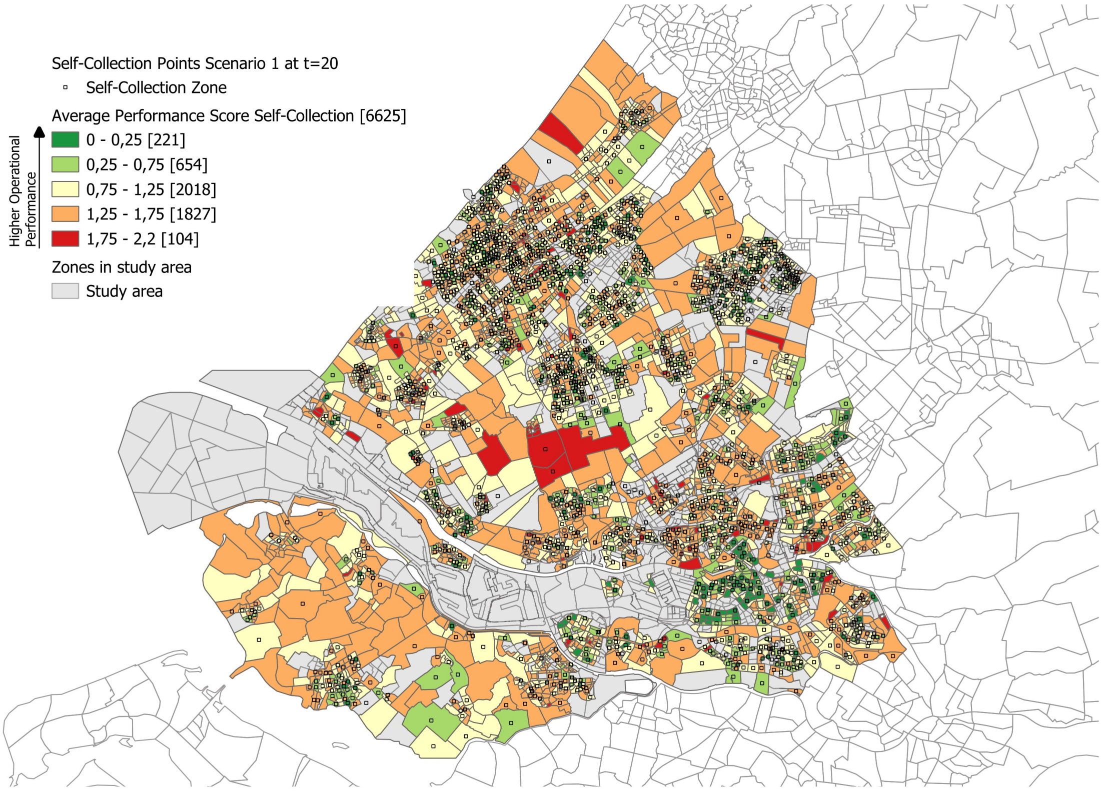
Figure 5.24: Zonal performance score in run 1 in scenario 1

The performance of the delivery methods can also be interpreted spatially. In Figure 5.24, the difference between the performance scores of van and self-collection delivery are shown per zone averaged over all iterations and at time point $t = 3$. A negative indicator corresponds with van delivery performing better in a zone, and a positive indicator with self-collection performing in a superior way. In correspondence with the average performance score of self-collection being higher than that of van, the number of zones where van delivery performs better is significantly higher. Even in lots of zones with a self-collection point, van delivery outperforms self-collection. However, the situation at $t = 3$ shows that in most zones with a self-collection point, that delivery method is conjointly the well-performing delivery method. The reduced performance of self-collection results from the limited capacity, which has consequences on the reliability score.

The zones where self-collection performs well are generally small with a high population density; see Figure 5.25. In the city centres, self-collection establishes a higher performance, and most zones do possess a self-collection point. Contrary, large rural zones result in mediocre to poor operational performance. This is both a consequence of the longer distance to pick up a parcel and the capacity that falls short because a self-collection point services multiple zones. Van delivery performs relatively constant across all zones. That outcome results from van delivery not being dependent on geographical inputs in this model. The successful delivery rate, a constant across all zones, solely influences the operation performance.



(a) Average performance score van



(b) Average performance score self-collection

Figure 5.25: Average zonal performance score in run 1 in scenario 1

5.3.2. Scenario 2: Crowdshipping

In scenario 2, consumers can also choose crowdshipping as a delivery method. As described in subsubsection 4.5.4.1, it is assumed that 21% of the consumers will prefer crowdshipping at $t = 0$. Likewise scenario 1, five runs are performed with 21 iterations, and the resulting average preference probability of each delivery method can be seen in Figure 5.26. In scenario 2, the variation in preference probabilities is limited.

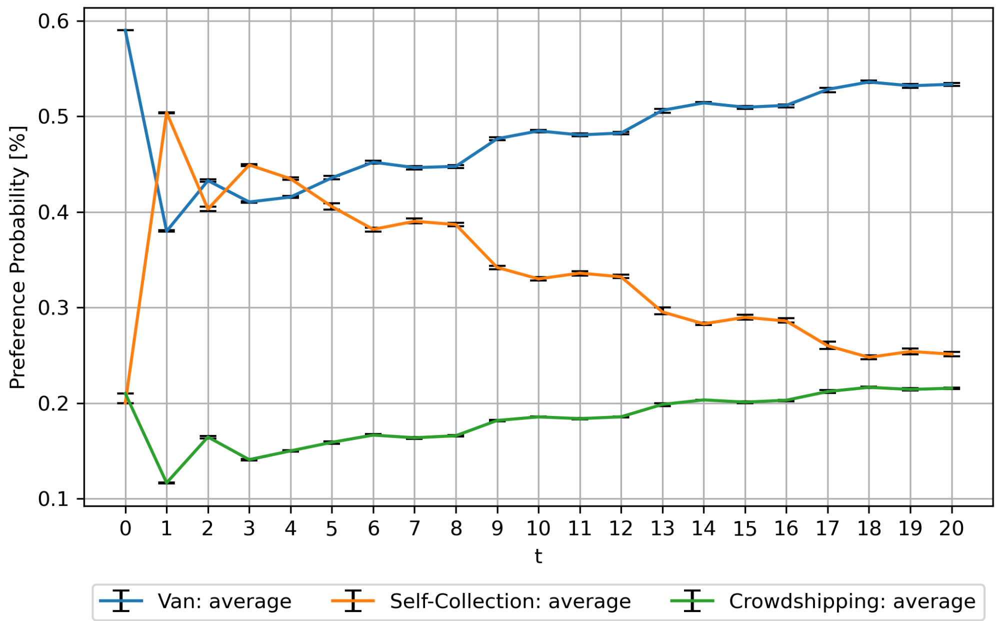


Figure 5.26: Preference probability in scenario 2

The preference probability of van and self-collection behaves similarly as in scenario 1. In the first four iterations, the demand for both delivery methods grows towards each other. The preference probability of crowdshipping halves in the first iteration. This mainly results from the performance score that is twice as high as that of the other delivery methods and this offering much lower service; see Figure 5.27. This performance score of crowdshipping is almost constant. Yet, the attribute scores, except the distance attribute, do evolve slightly; see Figure 5.28. However, those changes cancel each other out. The speed score rises marginally, and this is caused by the growing number of crowdship requests for which a crowdshipper is not found. This is also the reason for the slightly higher reliability score. As the speed level weights increase exponentially, un unmatched requests weights heavier than the growing number of requests for which a crowdshipper is found. The opposite is happing with the cost score due to the reduced compensation that is paid to a crowdshipper. The average compensation decreases from roughly €2,18 - 2,21 to €2,14 - 2,15. That arises from the growing crowdship parcel demand. As there are more requests, the chance of selecting a parcel request that best fits a traveller's journey increases, which results in lower compensation.

During the simulation, the self-collection performance score develops to the level of crowdshipping. This is purely because of the reliability score. A stepwise evolution can be seen. This effect occurs because the self-collection point distribution is adapted on a yearly bases, thus each four iterations. A redistribution results in a lower performance score of self-collection, consequently changing the preference probabilities.

As both van and crowdshipping have balanced operational performances, the probability of self-collection shrinks. Eventually, van stabilizes at roughly 53% preference probability, self-collection at 25% and crowdshipping at 21%. Due to the yearly parcel demand growth, the daily number of deliveries and trips that need to be performed increase strongly, mainly that of van delivery; see Figure 5.29.

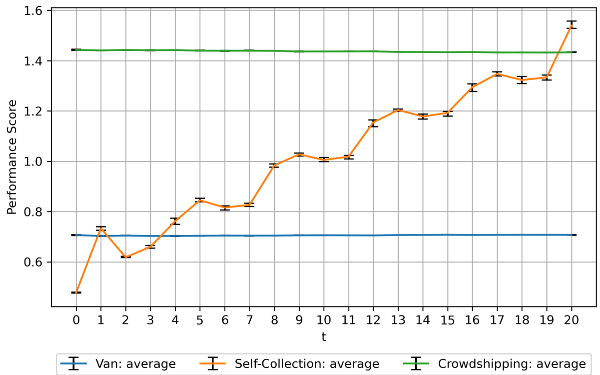
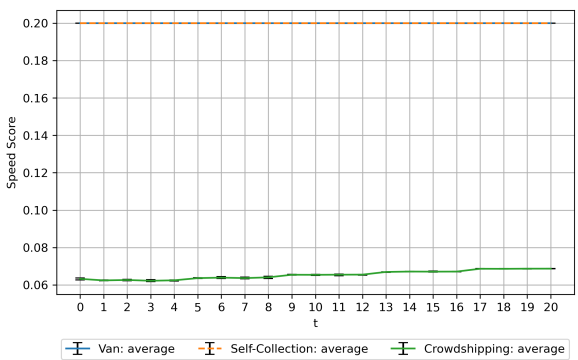
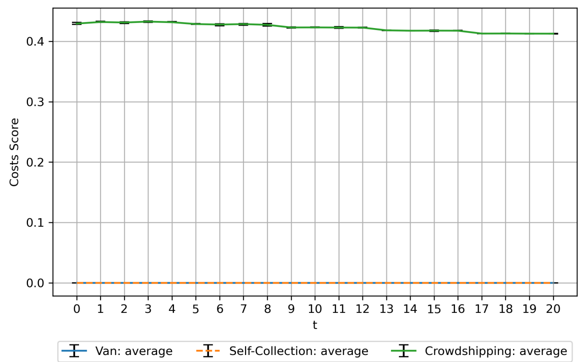


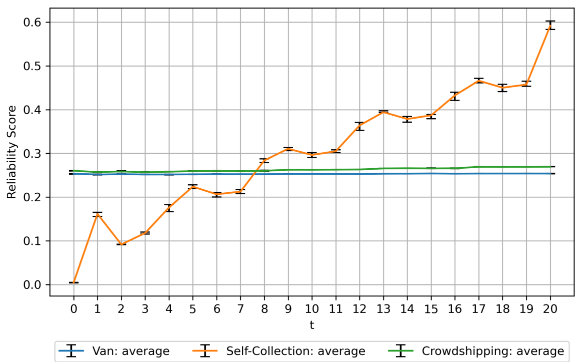
Figure 5.27: Performance score in scenario 2



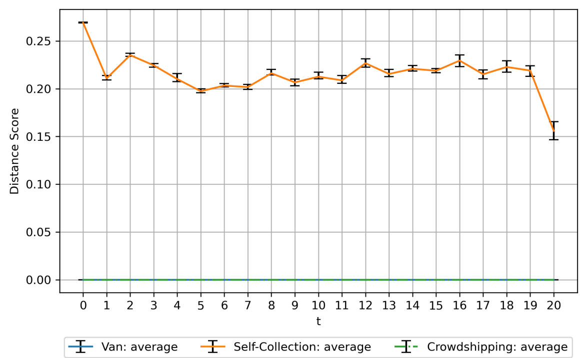
(a) Average speed score



(b) Average costs score

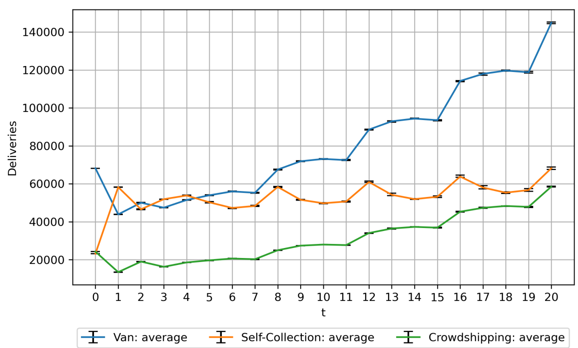


(c) Average reliability score

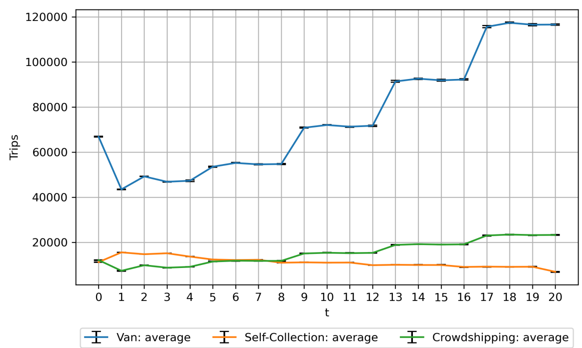


(d) Average distance score

Figure 5.28: Average attribute scores of all delivery methods in scenario 2



(a) Daily number of parcel requests



(b) Daily number of trips

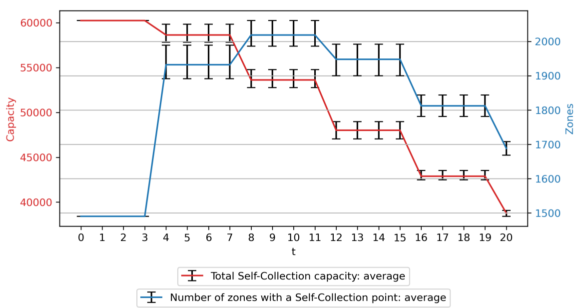
Figure 5.29: Daily parcel requests and trips in scenario 2

Table 5.11: Vehicle kilometres and emissions vans for zone to zone and intrazonal travel in scenario 2

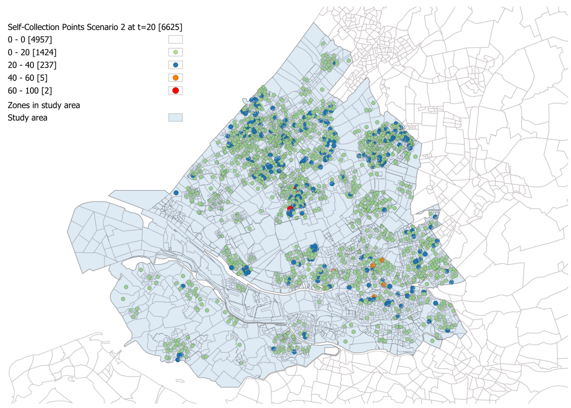
Travel	t=0				t=20			
	VKMs	VKMs %	kg CO2	CO2 %	VKMs	VKMs %	kg CO2	CO2 %
Zone-to-Zone	65403	79%	12333	74%	66723	69%	12680	64%
Intrazonal	17463	21%	4313	26%	29294	31%	7236	36%
Total	82866		16646		96017		19916	

Because of the lower demand for van delivery, the share of the zone-to-zone travel for vans is relatively larger. Similar to scenario 1, zone-to-zone travel is more important at the simulation horizon. This is, again, due to the effect of the fairly constant amount of zones that need to be visited and the increasing number of stops in each zone. With a growth of 92% for van and self-collection delivery, the vehicle kilometres of vans increase by 15.9%. Because of this, the emissions per parcel reduce significantly. For crowdshipping, the shares of trips performed by bike and car are equal. However, the total detour distance made by bikes is 15124 km and by cars is 30466 km at $t = 20$. This corresponds with CO2 emissions due to crowdshipping of 3047 kg at $t = 20$. In total, 141607 vehicle kilometres need to be made in scenario 2, resulting in a total CO2 emission of 22963 kg CO2.

In this scenario, the reliability score of self-collection increases, and similar to scenario 1, the service quality decreases. This is caused by the strong decline of the capacity of the self-collection points; see Figure 5.30a. After the first four iterations, the number of zones with a self-collection increases significantly. This follows from the high self-collection delivery demand in year one. However, as the demand stabilised in numerous zones, the self-collection points were removed. Eventually, this results in the distribution at $t = 20$ as shown in Figure 5.30b. In comparison with the initial distribution, see Figure 5.4, the trend in this scenario is removing self-collection points from a zone with multiple points to a zone without. Unlike scenario 1, the most urbanised areas, for example, zones in cities such as The Hague and Delft, are not yet saturated. Therefore, in this scenario, self-collection points are not distributed to less urbanised and rural zones.

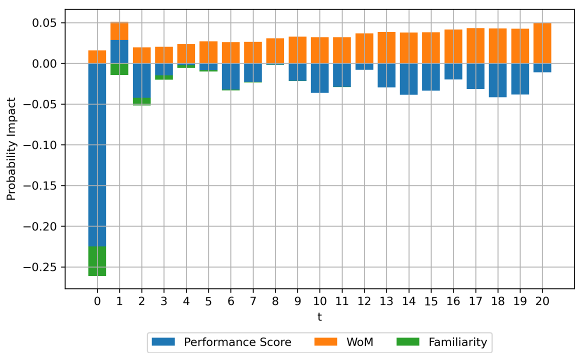


(a) Evolution of capacity and distribution self-collection points

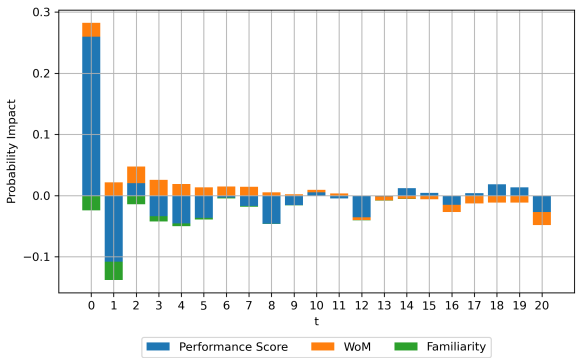


(b) Self-collection point distribution across the study area at $t = 20$

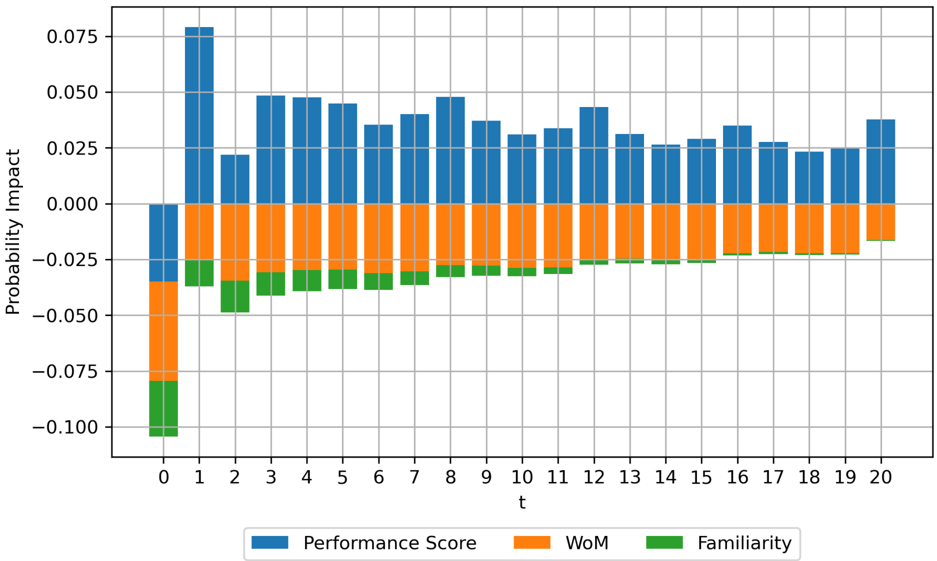
Figure 5.30: Evolution of self-collection points in run 1 in scenario 2



(a) Impact on the preference probability for van



(b) Impact on the preference probability for self-collection



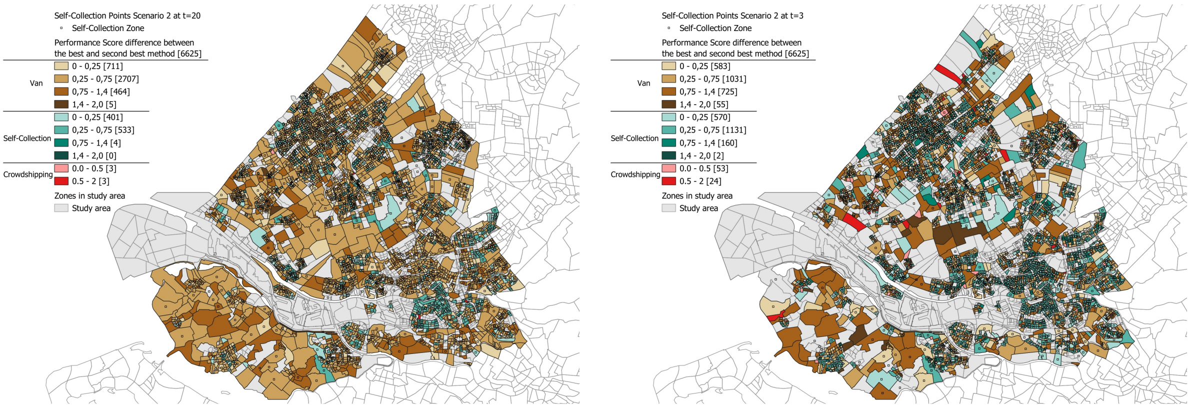
(c) Impact on the preference probability for crowdshipping

Figure 5.31: Preference probability impact in run 1 in scenario 2

In Figure 5.31, the impact on the preference probability of the performance score, WoM and familiarity are plotted for each delivery method. In the case of van delivery, the initial probability, although lower than in scenario 1, is still above the probability the logit model estimates. Therefore, a strong negative effect of the performance score occurs at $t = 0$. As the performance score of van delivery is below average, indicating a good operation performance, the WoM effect is always positive, even at $t = 0$. Because of the rising performance score of self-collection delivery, van delivery is relatively becoming a better delivery option. Therefore, the magnitude of WoM keeps increasing.

Self-collection shows strong growth at $t = 0$ because the performance score is the lowest, and thus the logit model estimates the highest probability. Likewise, this results in a positive WoM at the start of the simulation. From $t = 12$, the performance score inclines to above average, leading to a negative WoM effect.

Crowdshipping shows the least change in probability during the simulation. This is reflected by the magnitude of the impact on the probability that is lower than that of van and self-collection. In virtue of that, the familiarity effect can be clearly seen in crowdshipping. The constant negative effect narrows down over time, but in this case, it significantly impacts the preference probability in the first half of the simulation. The performance score and WoM, except at $t = 0$ oppose each other, and the WoM effect slows down the preference probability growth expected by the performance score.



(a) Average performance score difference

(b) Performance score difference at $t = 3$

Figure 5.32: Zonal performance score in run 1 in scenario 2

Van delivery has the highest average performance score over all iterations in most zone; see Figure 5.32a. For the greater share of those zones, the difference in performance scores is small to moderate. The biggest differences occur in large zones that generally do not possess a self-collection point. Self-collection has the best performance in various zones. However, there are no zones where self-collection is better to an extreme extent. Similar to scenario 1, self-collection performs well in highly urbanised and small zones; see Figure E.1. As can be seen in Figure 5.32b, self-collection has a higher quality operation at the start of the simulation. This difference results from the capacity reduction over time, which consequently reduces the reliability of self-collection.

On average, there are just six zones where crowdshipping performs best. Specific characteristics cannot describe those zones. This corresponds with the high-performance score of crowdshipping across all zones; see Figure 5.33. At the end of year one, $t = 3$, there are more zones in which crowdshipping comes on top. Typically those zones have a low demand in iteration 3, and van or sell-collection delivery had a failed delivery. Because of the limited number of parcels, those failed deliveries worsened the performance score of van and self-collection, consequently making crowdshipping the dominant delivery method.

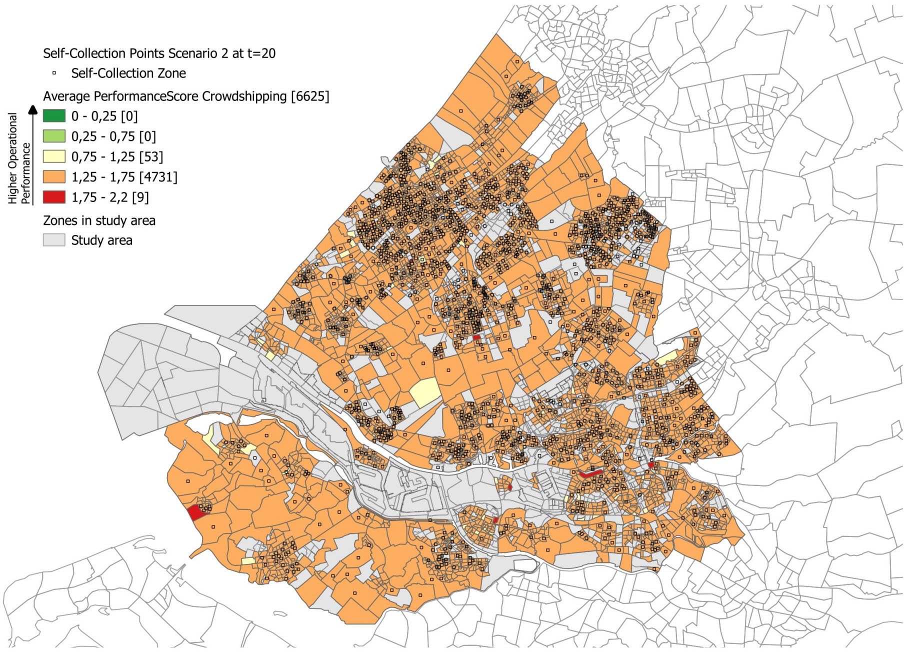


Figure 5.33: Average performance score crowdshipping in scenario 2

5.3.3. Scenario 3: Drone

Scenario 3 explores the introduction of a new delivery method: drones. Also for this new delivery method, the system capacity is a dynamic characteristic of the service. Thus at $t = 0$, consumers have not yet used that delivery method and the preference probability is 0%. Again, five runs are performed with each 21 iterations to model a five-year time horizon. The average preference probability of each delivery method can be seen in Figure 5.35. The minimum and maximum values are close to the average values, and along with the other scenarios, the variation of the preference probabilities is limited.

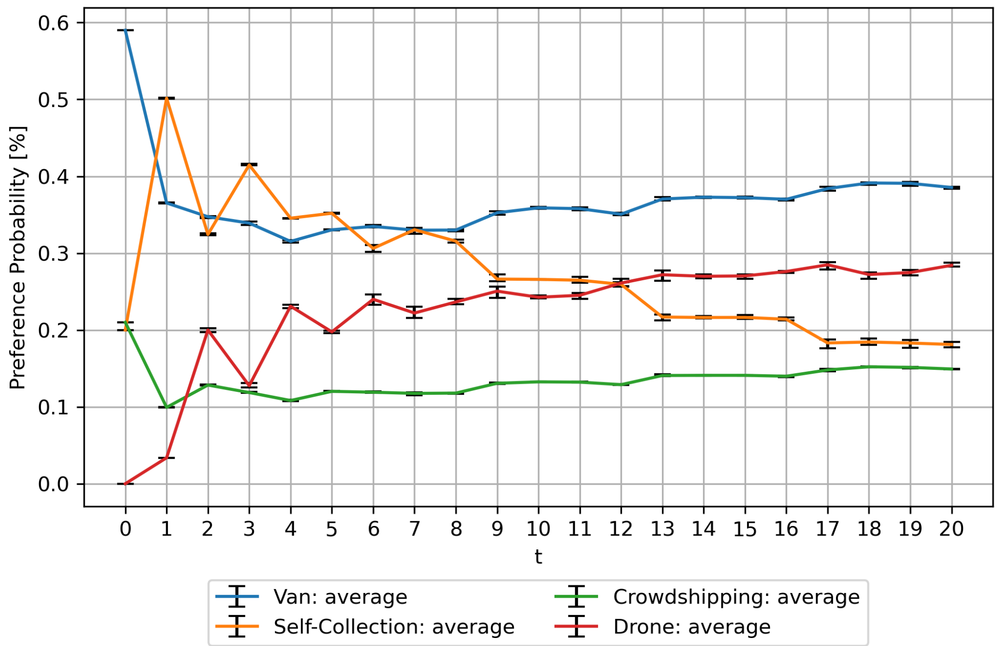


Figure 5.34: Preference probability in scenario 3

Van, self-collection and crowdshipping show similar behaviour as in scenario 2, although, certainly, at the end of the simulation, those delivery methods all lose market share to drone delivery. After the first iteration, non of the consumers used drone delivery. Because of the potential performance of drones, innovators try the service in $t = 1$. Consequently, in that iteration, a performance score for drone delivery can be calculated for the first time. The performance score is relatively high, and the service is thus not very satisfactory due to the high speed and reliability score; see Figure 5.36. Those high values are the consequence of the low capacity of drones. At $t = 2$, the number of drones at the depots is adapted based on the demand, which results in a growth from 1 drone at each depot to almost 200 drones in total; see Figure 5.37b. This results in a steep drop in the speed and reliability score, leading to a high level of service. From $t = 5$, the performance score of drone stabilises.

In Figure 5.35, it can be seen that the performance scores of self-collection and drone change over time and cross each other. Similar to the previous scenarios, the capacity of self-collection is strongly reduced, see Figure 5.37a, consequently worsening the reliability score.

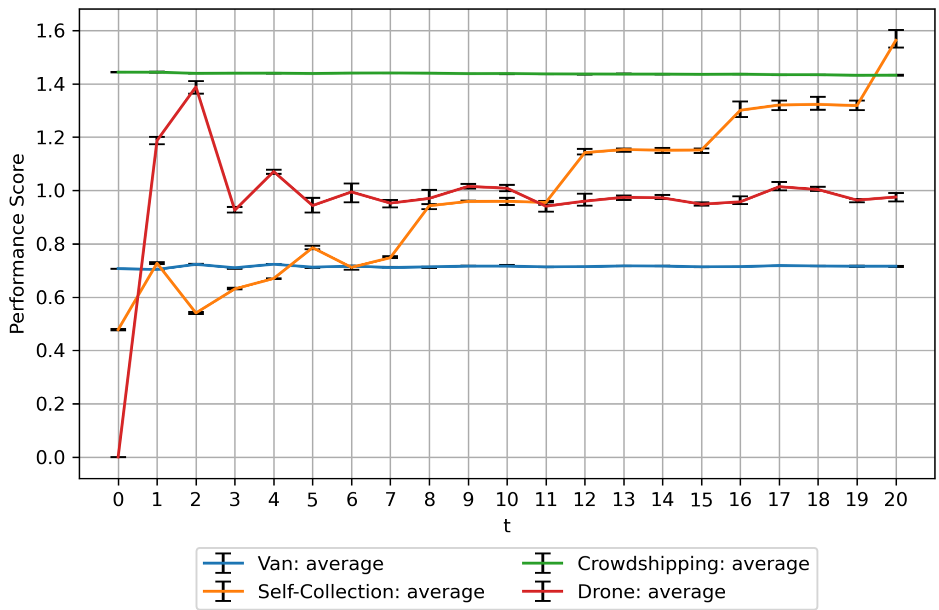
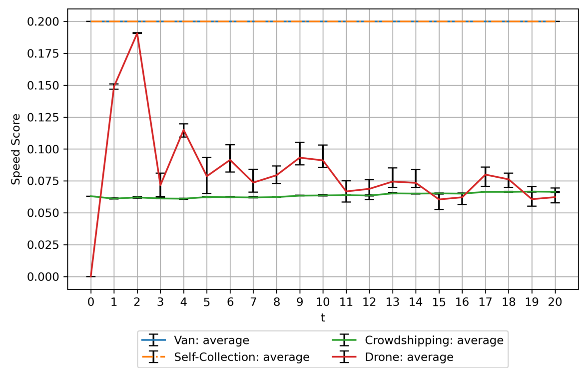
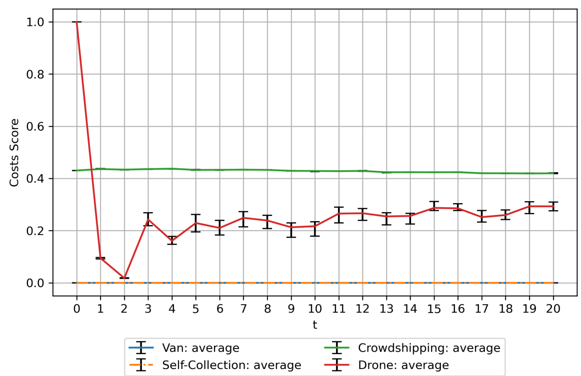


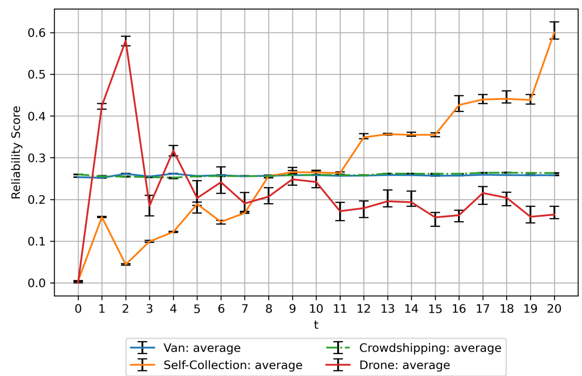
Figure 5.35: Performance score in scenario 3



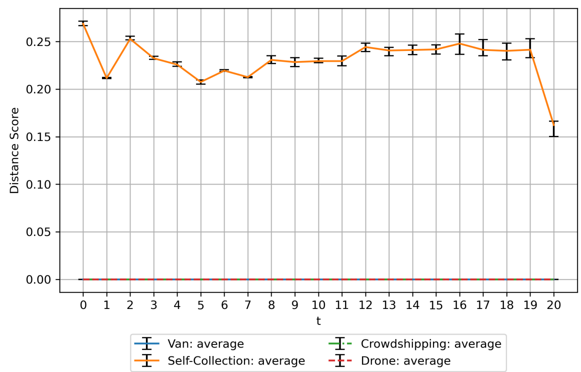
(a) Average speed score



(b) Average costs score

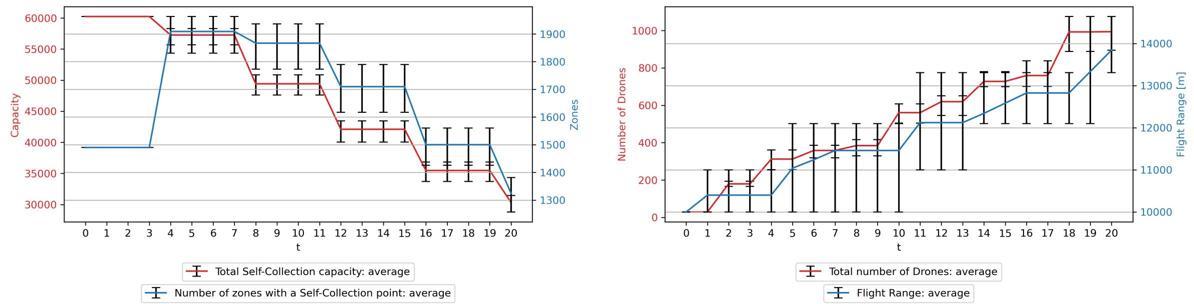


(c) Average reliability score



(d) Average distance score

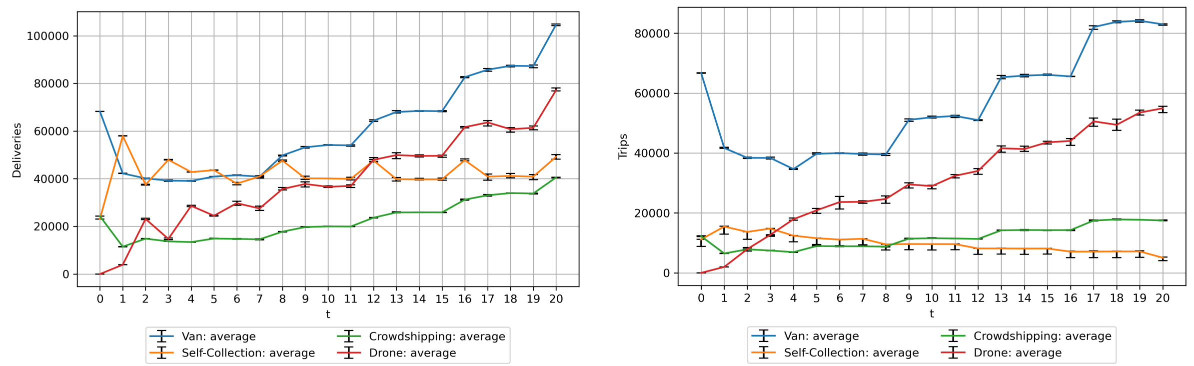
Figure 5.36: Average attribute scores of all delivery methods in scenario 3



(a) Evolution of capacity and distribution self-collection points

(b) Evolution of the total number of drones and the flight range of drones

Figure 5.37: Evolution of self-collection and drone operations in scenario 3



(a) Daily number of parcel requests

(b) Daily number of trips

Figure 5.38: Daily parcel requests and trips in scenario 3

Although the preference probability of self-collection delivery shrinks, the number of parcels that need to be delivered per day stays quite constant; see Figure 5.38a. The parcel demand for the other delivery methods increases significantly during the simulation. Especially the number of drone deliveries grows enormously, see Figure 5.38b. At the time horizon, van delivery has an estimated market share of 39%, and self-collection ends with a market share of 18%. Crowdshipping has the lowest share with 15%, for which cars ride just above 24600 additional kilometres and bikes 12000 km. Drone deliveries grow to 28%, for which roughly 36.000 parcels can be delivered by drone per day.

The vehicle kilometres of vans are very similar to scenario 2 at $t = 0$ because drone delivery is not used at the simulation start; see Table 5.12. However, at $t = 20$, the vehicle kilometres and emissions of van delivery are hardly higher, while the demand for van and self-collection delivery has grown by around 40%. The introduction of drones results in a demand for van delivery for which fewer zones need to be visited. Thus van delivery becomes more efficient in the number of kilometres per parcel, which lowers from 0.90 km/parcel to 0.65 km/parcel. For crowdshipping, 36785 vehicle kilometres are made at $t = 20$, and crowdshipping by car results in the emission of 2462 kg CO2. Drone delivery does not emit CO2 as they operate on electricity. However, the energy consumption per day is rather high, as 195858 vehicle kilometres are made. This large distance is the consequence of drones only being able to deliver one parcel per flight.

Table 5.12: Vehicle kilometres and emissions vans for zone to zone and intrazonal travel in scenario 3

Travel	**t=0**				**t=20**			
	VKMs	VKMs %	kg CO2	CO2 %	VKMs	VKMs %	kg CO2	CO2 %
Zone-to-Zone	64889	79%	12238	74%	62816	75%	11904	70%
Intrazonal	17384	21%	4294	26%	20389	25%	5036	30%
Total	82273		16531		83205		16940	

Next to the great number of vehicle kilometres that drones need to make, drone operations will be very demanding. In just five years, the drone fleet has to grow from 29 drones in total to almost 1000 in order to fulfil the demand. Consequently, there will be four depots that must operate more than 100 drones, with a maximum of 183. As one drone delivery takes roughly 20 minutes on average, the operational pressure on a depot is immense. It is questionable if these model results could be achieved in real-world operations within a time frame of five years.

The innovation of drones does improve the operation of that delivery method. As the flight range increases, the reliability score improves (the reliability becomes lower) as more zones can be reached from the depots. That effect is marginally countered by the rise in the delivery costs, which is linked to the travel distance, and thus the flight range. As a consequence, the demand for drone delivery only increases slightly because of the improved capabilities of drones.

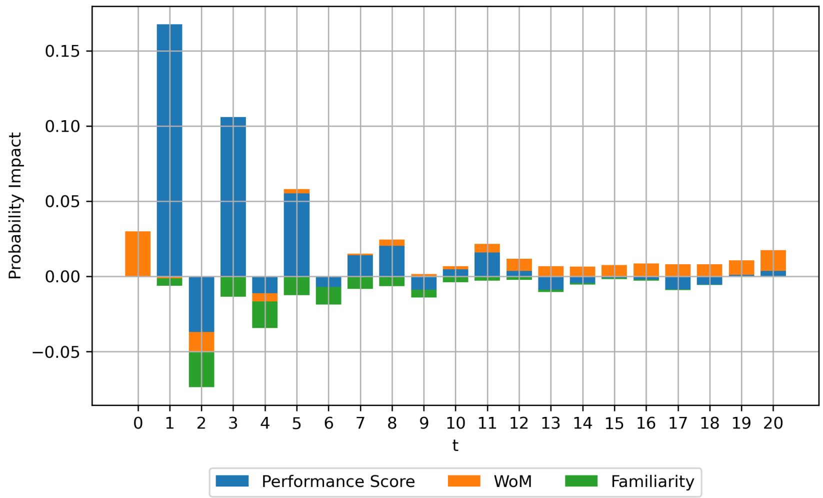


Figure 5.39: Preference probability impact for drone in run 1 in scenario 3

The impact on the preference probability of van, self-collection and crowdshipping develop in the same way as in scenario 2 and can be seen in Figure E.2. The preference probability impact of drone delivery is shown in Figure 5.39, and at $t = 0$, the only influence comes from the WoM effect. This is because that factor accounts for the adoption of innovators. Based on the operational performance that those innovators experience, the logit model assumes a strong uptake in users at $t = 1$. At that iteration, the magnitude of WoM is very limited due to the small group of consumers that can spread WoM. The familiarity effect is relatively strong at the start because only a few consumers have used drones. When the performance score of drone delivery stabilizes, also the impact becomes smaller. In the end, the WoM effect becomes the most important reason for the continued growth of the market share of drones.

In scenario 3, the demand for self-collection is the lowest, and consequently, this leads to a more sparse distribution of the self-collection points; see Figure 5.40. The number of points slinks, resulting in fewer zones with multiple points as well as fewer zones with at least one self-collection point. Both in urbanised cities and more rural villages, points are removed. However, as the supply of points in less urbanised areas often consisted of only one or a few points per municipality, in this scenario, the number of points becomes quite scares outside large cities. Correspondingly, the average performance score over the zones shows a clear distinction between urbanised and rural zones; see Figure 5.42. Both the performances of van and crowdshipping across the study area are comparable with previous scenarios.

In Figure 5.41a, the average performance score of drone delivery of each zone is shown. As drones can fly within a range from their starting depot, it can be noticed that generally, the performance scores of zones at the edge of the study area, thus further away from most of the depots, is worse; see Figure 5.41b. This is a result of drones not being able to operate in those zones. As many of those zones are also in more rural areas, the zones where drone operations have a lower performance are also the

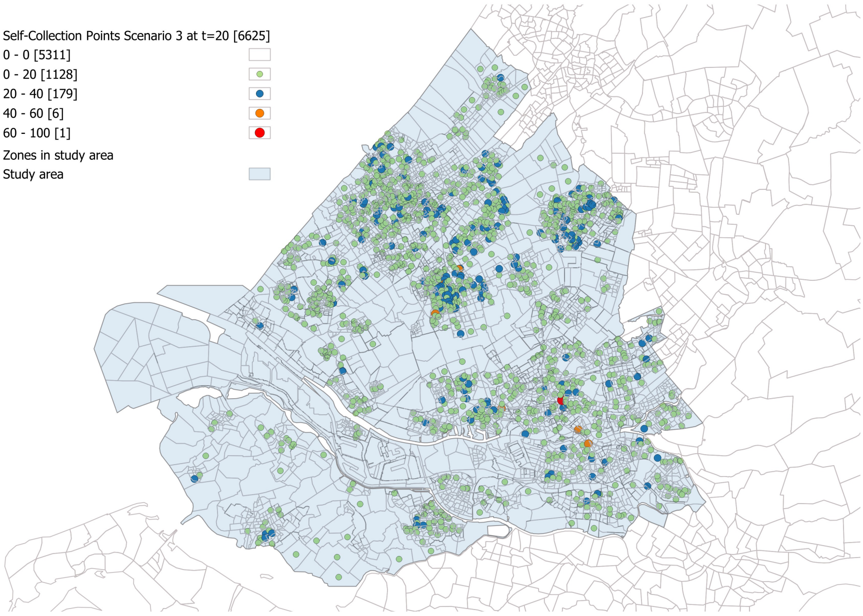
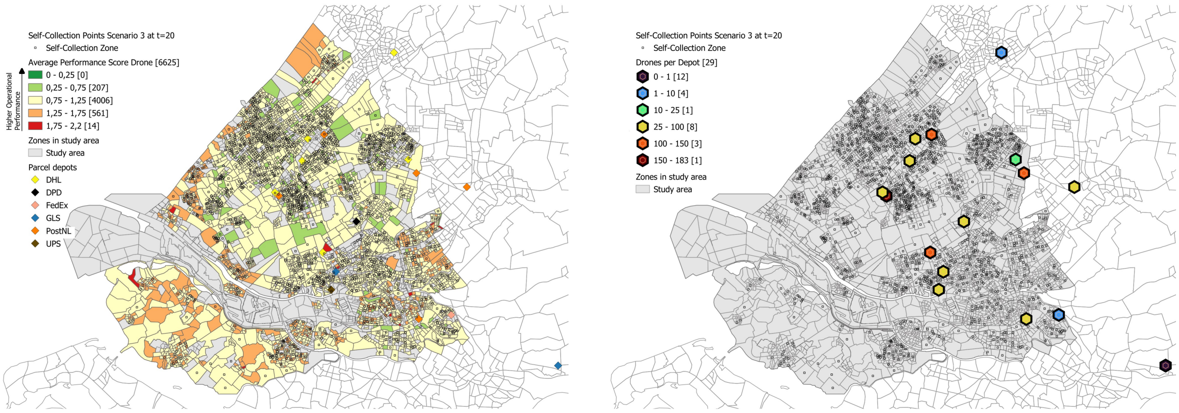


Figure 5.40: Self-collection point distribution across the study area at $t = 20$ in scenario 3



(a) Average performance score drone in scenario 3

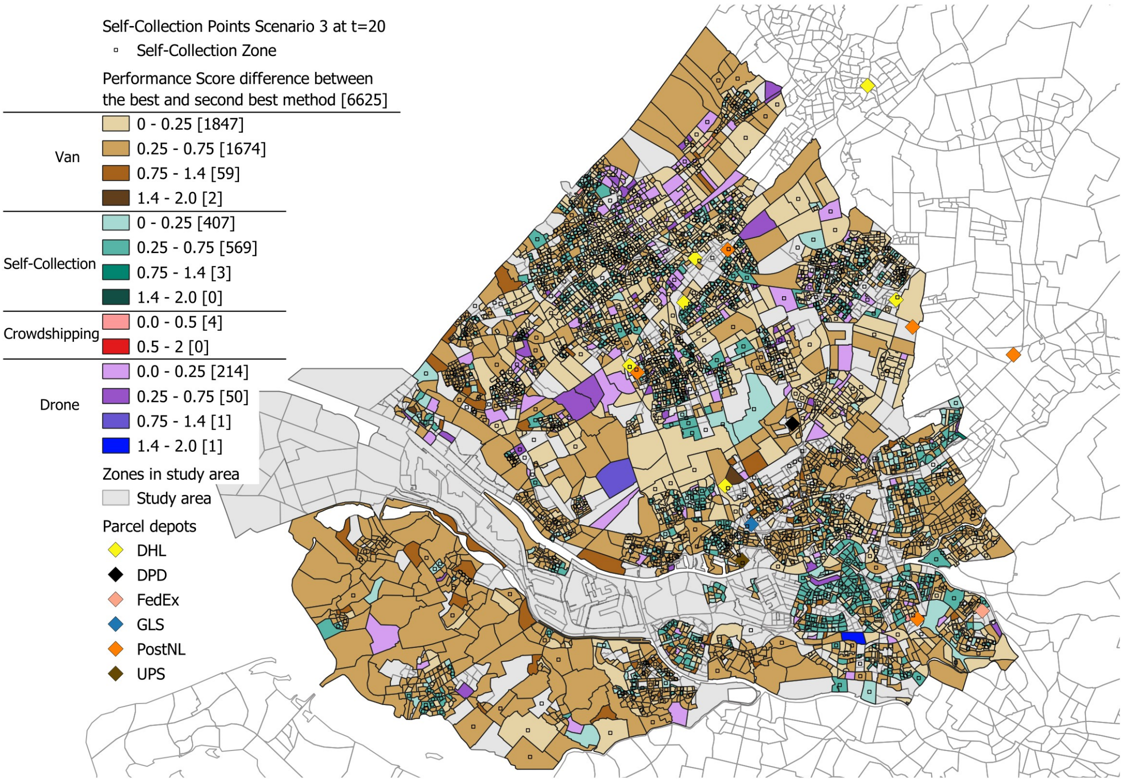
(b) Drones per depot at $t = 3$

Figure 5.41: Drone performance in run 1 in scenario 3

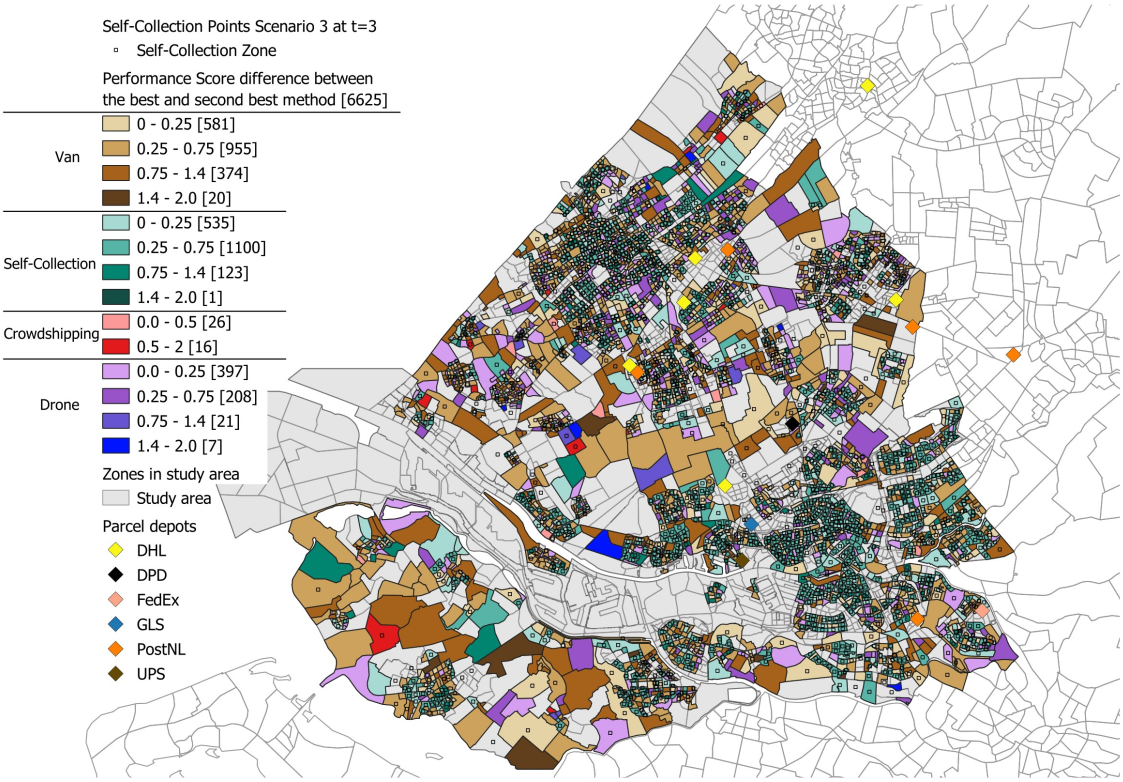
zones where self-collection performance is porer.

The introduction of drones as a delivery method results in a partially different zonal distribution of best-performing delivery methods. In the centre of the study area, drone delivery is often a strong addition that outperforms the earlier used delivery methods. However, in city centres, self-collection maintains an outstanding operation. Still, van delivery is dominantly the highest-performing delivery method. Certainly, in less urbanised zones, especially those on the outskirts of the study area, van delivery achieves better scores. Crowdshipping has a very limited value in being the most excellent performing delivery method in this scenario.

At $t = 3$, see Figure 5.42b, the results are more varied. Particularly, self-collection is more prevalent in this stage. The delayed decline of self-collection is the consequence of the reduced number of self-collection points. However, also drone delivery outperforms the other delivery methods in more diverse zones. This result forms because, in many zones, only one drone request is made at $t = 3$. Thus the chance of something going wrong with the delivery is lower than that of other delivery methods, as those have to perform multiple deliveries in most zones.



(a) Average performance score difference



(b) Performance score difference at $t = 3$

Figure 5.42: Zonal performance score in run 1 in scenario 3

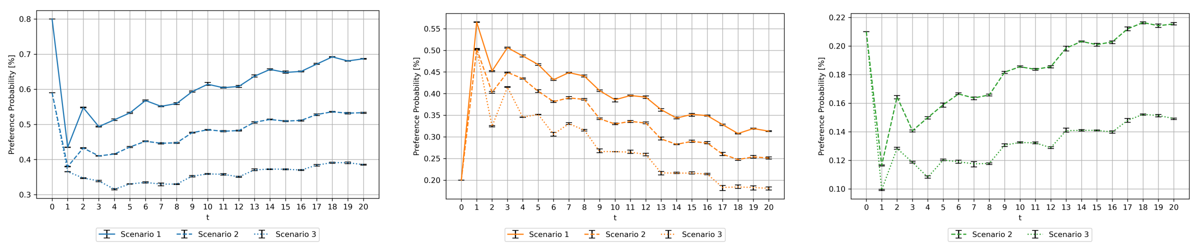
5.3.4. Scenario Comparison

The results indicate that most consumers prefer van delivery in all three scenarios. The performance score of van delivery is constantly low, which represents a high operational service quality. This is mainly a consequence of the constant reliability of vans. In the iterations near the end of the time horizon, the WoM effect pushes consumers to choose this option. The growth in parcel demand leads to an enormous increase in the number of trips that have to be performed per day; see Table 5.13. The addition of other delivery methods can significantly reduce the demand for van delivery, see Figure 5.43. Although the performance scores of the other delivery methods are, on average higher, thus performing worse, many consumers will choose those delivery methods.

Table 5.13: KPIs scenarios

Indicator	Method	Scenario at t=0			Scenario at t=20		
		1	2	3	1	2	3
Market share	Van	80%	59%	59%	69%	53%	39%
	Self-Collect	20%	20%	20%	31%	25%	18%
	Crowdshipping		21%	21%		22%	15%
	Drone			0%			28%
Trips per day	Van	92485	66815	66716	153277	116497	84181
	Self-Collect	11343	11171	11171	8052	6890	5053
	Crowdshipping		12079	12237		23169	17729
	Drone			0			36324
Vehicle kilometers per day	Van	90394	82866	82273	108793	96017	83205
	Crowdshipping		26492	26702		45590	36785
	Drone			0			195858
CO2 emissions per day	Van	18484	16646	16531	22904	19916	16940
	Crowdshipping		1782	1792		3047	2462
	Drone			0			0
Number of Self-Collection points		3012	3012	3012	2250	1956	1508
Total capacity Self-Collection points		60240	60240	60240	44500	39000	30000
Number of Drones				29			994

The estimated shares for self-collection decline over time in all scenarios because the supply of self-collection points becomes smaller. Still, around one-fifth of the consumers are expected to choose self-collection. Because self-collection delivery allows for grouping parcels, it is possible to deliver more parcels in fewer trips. The spatial results show that self-collection can be a competitive delivery method in dense urban areas. In scenarios 2 and 3, the results show that fewer consumers will prefer self-collection. Because of those lower demands, the distribution of self-collection becomes scarcer, and the total capacity at $t = 20$ is around 13% and 31% lower for scenarios 2 and 3, respectively. Consequently, the performance score of scenario 3 evolves to a slightly worse level than in scenarios 1 and 2.



(a) Preference probability Van for the three scenarios

(b) Preference probability Self-Collection for the three scenarios

(c) Preference probability Crowdshipping for the three scenarios

Figure 5.43: Preference probabilities for the three scenarios

Crowdshipping has, like van delivery, a constant performance score over all iterations and in both scenarios 2 and 3. The operational performance of crowdshipping is the poorest until the last iteration when self-collection becomes worse. Because of the constant performance, the evolution of crowdshipping is very comparable between scenarios 2 and 3. However, in scenario 3, the competition with drone delivery results in a lower market share.

The introduction of drone delivery has a large impact on the predicted market shares of the other delivery methods. Drones do perform well around the depots, for which almost a thousand drones that operate 18 hours a day are needed to fulfil the demand. However, the performance in the outskirts of the study area is generally low.

To conclude, the results show that parcel delivery will have a strong reliance on vans. Especially as that delivery method forms a backup delivery option for crowdshipping and drones. The performance score of the delivery methods do not differ strongly between the three scenarios and are relatively close to each other. Therefore, it is estimated that all delivery methods will end with a significant market share. The differences in the market shares primarily result from the addition of new delivery methods.

Finally, the large yearly growth in parcel demand will result in higher vehicle kilometres and CO2 emissions, regardless of the scenario. As can be seen in Table 5.14, the total demand almost doubled in five years. The introduction of crowdshipping reduces the demand for van delivery and self-collection, yet it will result in more vehicle kilometres that need to be made on the same road infrastructure as van delivery. However, crowdshipping produces less CO2 per km, and therefore the CO2 emissions are very comparable with scenario 1. The CO2 emissions in scenario 2 are still high because the positive effect of the substitution of van and self-collection delivery by crowdshipping is partly reversed. This is due to vans becoming more efficient when the demand is higher. With increasing parcel demand, the number of zones that need to be visited increases slightly, and that contributes to the largest share of vehicle kilometres and emissions. This can also be seen in Table 5.15, where the average kg CO2 emission per parcel is shown. In all scenarios, this number improves. This is mainly because van delivery has become more efficient. Scenario 3 results in the lowest CO2 emissions due to drone delivery being emission-free and results in 15% less CO2 emission than scenarios 1 and 2. Additionally, the van vehicle kilometres and emissions within zones are much lower than in the other scenarios, which reduces the last-mile burden in urbanised areas and on low-capacity roads. However, to operate that last-mile scenario, drones must fly an enormous amount of kilometres per day, and depots have to build large facilities to implement a drone delivery service of this quality.

Table 5.14: Scenario comparison

Indicator	Scenario at t=0			Scenario at t=20		
	1	2	3	1	2	3
Parcel Demand	115608	115608	115608	223318	223318	223318
Vehicle Kilometres	90394	109358	108974	108793	141607	315848
CO2 Emissions [kg]	18484	18429	18323	22904	22963	19402

Table 5.15: Average C02 emissions per parcel [kg/parcel]

Time	Scenario 1	Scenario 2	Scenario 3
t = 0	0.160	0.159	0.159
t = 20	0.103	0.103	0.087

6 Conclusion & Recommendations

6.1. Conclusion

This section will present the conclusion of this thesis. This study aimed to explore the interconnection between consumer preferences for parcel last-mile delivery methods and the evolution of the available delivery methods. First, the research sub-question will be discussed, and finally, the main question will be answered.

RQ1: What kind of parcel delivery methods are there or will likely be developed in the coming years?

Human-driven vans, cargo bikes, pick-up points, and automated lockers are well-established parcel delivery methods in the Netherlands. Drone and droid delivery are under development, and real-world tests are being executed, although not yet in the Netherlands. Also, crowdshipping and unattended home delivery innovations have the potential to alleviate the last-mile problem. In this thesis, it is chosen to implement van, self-collection, crowdshipping and drone delivery. The main reason for this choice is that all four delivery methods have distinctive characteristics that give the consumer a unique experience during the handover.

RQ2: What factors influence the last-mile delivery choice of consumers, and can consumers be grouped based on their preferences on those factors?

Consumers evolve their preferences based on multiple factors. In the first place, the diffusion of innovation theory presents that the perceived characteristics of a consumer cause certain adoption behaviour. Consumers use their experience, the perceived performance, to evaluate if their chosen delivery method provided them with a satisfactory service. Attributes can be used to express the operational performance of a delivery method. A literature review revealed that delivery speed and costs are the most widely used attributes, followed by time slots and parcel tracking. The last two are not applied in this study. A reliability attribute is incorporated that can describe failed deliveries. Furthermore, as only self-collection delivery requires consumers to pick up their parcel, an attribute is added for the distance consumers need to travel to pick up their parcel.

A second factor used to analyse consumer preferences dynamics is word of mouth (WoM), which describes the process of consumers evaluating a service and communicating that experience with other consumers. Especially for new services, consumers rely on the experiences and opinions of others in their decision process. An S-shaped adoption curve characterises the WoM effect. That effect arises from the growing group of people that can spread WoM due to their experience and a shrinking group of consumers that can still adopt the service. The Bass diffusion model is a regularly used formulation to analyse this WoM effect on market potential.

Familiarity is another factor that describes the dynamics of consumer adoption of services. According to this theory, consumers build trust and loyalty towards earlier used services or products, making them less likely to opt for an unknown delivery method. Research shows that consumers have difficulty predicting the performance of those unfamiliar services and that naive consumers are, to a smaller extent, able to select attributes that are predictive of the service performance.

Literature indicates that e-consumers are a heterogeneous group. However, a definitive consensus on which socioeconomic and demographic data cause these groups has not been developed. The studies of Caspersen and Navrud (2021), Buldeo Rai et al. (2021) and Pani et al. (2020) all present that certain groups are more open to using new delivery methods because those methods provide a better performance, are more environmentally friendly or fit closer to a particular lifestyle. Considering that most studies did not find significant socioeconomic and demographic indicators for such groups, it is chosen not to model distinctive consumer groups.

RQ3: Which dynamic feedback loops quantify the connection between delivery services and user preferences?

Consumers base their preferences for parcel delivery methods on various inputs, which often depend on the operational performance of the delivery methods. Based on literature a causal loop diagram is constructed to present the feedback loops in the parcel and consumer environment. This diagram forms the basis for the simulation model.

Four feedback loops can be identified. The *Reliability Evaluation* loop represents the overreaching of the capacity of a delivery method, resulting in a devalued reliability, creating a balancing feedback loop. To this overreach, the system will react, with a delay, via the *Capacity Evolution* loop, where the capacity of that delivery method can be modified to match the demand. The *Word of Mouth* loop represents a mental mediation factor, where the preference for a good-performing delivery method will be boosted due to the WoM effect. An important exogenous variable is the *Parcel Demand Growth Rate*, as this strong growth puts high pressure on the last-mile. Furthermore, the familiarity variable influences the *Delivery Preference Evaluation*. At that auxiliary variable, all input of the delivery methods is used to estimate the new preference probabilities. The Diffusion of Innovation theory is applied, which suggests that preferences directly affect people's behaviour. Thus, the preference probability results in the choice of a delivery method and a corresponding demand.

The *Distance Evolution* loop is specific for self-collection delivery and describes that with high demand, the number of self-collection points can be increased with a delay. Likewise, the *Innovation of Method i* is only applicable to drones, and it simulates that the performance of this developing method can be improved during the simulation time. The *Delivery Costs Method i* and *Delivery Speed Method i* are defined per delivery method and are calculated for each parcel delivery. Based on the CDL, it can be concluded that a high operational performance will lead to a high demand. Consequently, a dynamic interaction will develop where the delivery method has to deal with the high demand.

With the estimated market shares of the delivery method, several other KPIs can be derived. The demand for each delivery method is linked to the trips and corresponding emissions of those delivery methods. Increasing the number of trips does increase the emissions. Parcels for self-collection will be combined with the transport for van delivery.

RQ4: What is an effective modelling approach to simulate the dynamics between consumer preferences and the delivery service?

Because of the various feedback loops that interact in the CLD the System Dynamics methodology is selected to model the system described in the CLD. An essential concept of SD is that feedback and delay cause the behaviour of a system. With SD a system is replicated at the aggregate level, which brings low data requirements. Additionally, it can visualise the interdependencies between different system parts and modelling algorithms from various fields can be integrated.

The SD model is combined with an agent-based urban freight transport model: MASS-GT. The urban freight model simulates the grouping and delivery of parcels and estimates the corresponding routes and emissions. Additionally, the results from MASS-GT provide input to determine the operational performance of each delivery method. With that operational performance, the SD model estimates the preference probability in the next iteration. By integrating the ABM and SD model, the dynamics between consumer preferences and operational performance are simulated.

The developed simulation model is verified and validated with multiple tests, and it is concluded that the model results can be used for the purpose of this study. A calibration and several sensitivity analyses are conducted to improve the model results. It is established that the model is most sensitive to the beta weights of costs and reliability, that the impact of WoM and familiarity are moderate, and that the improvement of drones due to innovations does only marginally affect the total performance of drone delivery. The evolution thresholds of self-collection, especially the threshold for removal, strongly impact the distribution and, thereby, the performance of self-collection. In the case of drones, that threshold does not result in significantly different performance of drone delivery.

RQ5: To what extent do consumer preferences for last-mile delivery evolve over time under different service scenarios with their system capacity and price-setting?

The model results show that the preferences of consumers differ per scenario and per time point in the simulation. The addition of delivery methods severely changes the market shares of the common and well-known van and self-collection delivery. Because both crowdshipping and drone delivery provide a service that is close to that of van and self-collection, it is estimated that the new delivery methods are preferred by 15% and 28% of the consumers, respectively, in the scenario where all four delivery methods are available. Both for van and self-collection, this means a reduction in the market share of more than 40%.

Furthermore, it can be concluded that consumers evolve their preferences over time. The performance scores of self-collection and drone delivery depend strongly on the demand and the capacity that the service can offer at that point in time. For both delivery methods, a dynamic reaction can be seen between the consumer demand and the altering supply side. In the case of self-collection, the model predicts that fewer self-collection points are needed. However, that reduction in supply leads to poorer performance, because of which fewer consumers will choose self-collection. Thus a negative reinforcing loop originates. An opposite effect happens with drone delivery, where the number of drones greatly increases because of the demand. With that reaction of the supply side, drones can guarantee a high level of performance at later time points.

In addition, the spatial results show that there are zonal differences in the performance of the delivery methods. Self-collection can achieve high operational performance in dense and urbanised zones. The results indicate that it is a sensible delivery method in large cities and the centres of villages. The drone service is strongest in the vicinity of the carrier depots. A strong level of service is realised in the central part of the study area, where the depots are mostly located. Van and crowdshipping performances are quite insensitive towards the location. As both self-collection and drone delivery is less occupied for zones at the edge of the study area, for those zones, van delivery seems the best option from the consumer perspective.

The results confirm that the effect of performance, WoM and familiarity influence the preferences of consumers. The performance score has a strong impact, certainly at the start of the simulation. It pulls the preference probabilities towards the estimates from the logit model. The WoM effect is regularly the most dominant factor at the time horizon. Thereby, the direction often opposes that of the performance score. This indicates that consumers could deviate from a preference because of the experience of others. The familiarity effect yields the least impact of the three factors and logically has the strongest influence on delivery methods with low usage.

Main Question: What is the impact of evolving consumer preferences for last-mile delivery on parcel freight logistics?
First of all, the vast growth in parcel demand puts a heavy burden on last-mile logistics, especially on van delivery. In a scenario with only van and self-collection delivery, the vehicle kilometres will increase by 20% and the CO2 emissions by 24%. The CO2 emission will grow relatively faster because the additional vehicle kilometres are mostly intrazonal trips, thus on less efficient city roads. Although crowdshipping emits less CO2 per kilometre than vans, it will not automatically reduce the total CO2 emissions of the last-mile. Van delivery becomes more efficient when the density of parcels increases. Thus, a decline in demand causes a very limited reduction in vehicle kilometres and CO2 emissions. Yet, model results show that when a large share of the consumers shifts to low-emission delivery methods, it is possible to limit the growth in CO2 emission.

This case study indicates that from a consumer perspective, van delivery provides the best service in general. However, other delivery services can introduce competitive services, especially in highly urbanised zones. Because the operational performances of van, self-collection, crowdshipping and drone delivery are relatively comparable, all delivery methods gain a significant market share. Nevertheless, the market shares heavily depend on the service a delivery method can offer. Consumers are sensitive to the costs and reliability of last-mile services. The expected delivery costs, or fees, are very constant in this study. Accordingly, fluctuations in reliability play an important role in consumers' decisions. The reliability attribute is not only a crucial aspect in the evaluation of an experienced delivery, the performance score, but also in the WoM effect. Furthermore, even for new delivery methods, the impact of familiarity is limited compared with that of the performance score and WoM. This indicates that consumers are willing to choose an unknown delivery service if they expect improved performance.

The introduction of an extra delivery method does disrupt the prevailing situation. Consumers will try that new delivery method, thereby the adaptiveness of the delivery methods to the new demand is essential. In all three scenarios, at the end of the simulation, the market shares stabilize. This suggests that when the operations of last-mile delivery services maintain a constant level, consumers will not significantly change their preferences.

To conclude, the results of this study reveal that, regardless of the introduction of new delivery methods, the average number of kilometres by van per parcel and the average emission of CO2 per parcel will reduce with a growing parcel demand. This is because delivery by vans to consumers' homes and self-collection points becomes more efficient with the increasing density of the requests. However, that comes with the disadvantage that a larger share of the vehicle kilometres and emissions will be on intrazonal travel, bringing complex problems such as congestion and nuisance to urbanized areas. To reduce the total emissions and solve the congestion problem of van delivery, other delivery methods must be introduced. Self-collection delivery and drone delivery can offer high levels of service in many zones, and it can be expected that a large share of consumers will choose those delivery methods. Large investments will be needed to develop the self-collection network and especially the drone facilities to deal with that demand. The dynamics in this model show that if those investments are not made, a negative feedback loop will originate, because of which the consumer preference for self-collection and drone delivery will diminish.

6.2. Recommendations
This section discusses the scientific contribution of this research, followed by the limitations and potential improvements of the developed simulation model. Furthermore, recommendations for future research are presented, and the main takeaways for policymakers are elaborated.

6.2.1. Scientific Contribution
In this thesis, a last-mile delivery consumer preference simulation model is developed to capture and explore the interconnections between the performance of available delivery methods and consumer preferences for those delivery methods. By doing so, this study contributes to exploring the scientific gap in empirical studies into the dynamic interaction between consumer preferences and last-mile operation. An SD model and an ABM, respectively for consumer preferences evolution and last-mile

operation simulation, are combined to create accurate estimations of the operational performance of different last-mile delivery methods with diverse demand levels. This modelling approach provides insight into the direction and magnitude of various factors that take place, with the added novelty of gathering that empirical data at multiple time points in the study time horizon. This is a valuable model feature, as SP and especially RP research cannot determine such data.

Another contribution is the integration of multiple delivery methods in consumer-centred research. Such integration is scarce, consequently making it hard to predict the operational performance in a quantitive way. Furthermore, this model is combined with a traffic model, which links the demand and supply for parcel delivery with the resulting number of trips and emissions. In addition, model results show detailed zonal information in an extensive study area, thereby presenting a spatial outcome that can be used for socio-demographic research. Lastly, this research provides, to the writer's knowledge, a new, although simple, formulation to express the familiarity effect in quantitative preference probability changes.

6.2.2. Model Improvements

Although the developed model is able to describe the last-mile environment and the evolution of consumer preferences, and the results are verified and validated, various improvements could be implemented:

- First of all, the model is supported by many assumptions, like the beta weights, delivery costs for drones and the imitators coefficient in the WoM formula. A sensitivity analysis is performed for many of those assumptions. However, model results can differ significantly for sensitive parameters or extreme values. In the next section, Data Collection, suggestions are given to improve these assumptions or data sources. Nevertheless, it is not recommended to see this current model as an exact forecasting model.
- The consumer preferences are estimated at the aggregate level of the entire study area. The delivery performance differences between zones or municipalities are thus not considered. By applying the SD model at a lower level, for example, a neighbourhood or municipality, a more precise estimation can be given on where particular delivery methods prevail or are irrelevant. Implementing the SD model at a lower level is straightforward, as the SD model can almost directly be applied at a lower level. Additionally, the computation time of the SD model is very modest. Thus it is expected that performing the SD process a couple of thousand times increase the total computation time by a maximum of 10%. Thereby, a supplementary model step must combine all zonal information to the ABM model for the evolution of the delivery methods.
- In line with the previous point, by performing the preference evolution at the highest level, it is presupposed that all zones have similar beta weights for the delivery attributes. Differentiating these values can account for socio-demographic diversity.
- This SD method is chosen to model the consumer preference evolution, a reasonable approach in a system with various dynamic interconnections. However, this approach assumes homogenous decision-making. ABM is an alternative method that can simulate individual consumers, agents, and thus unique consumers. This is a well-established model approach to explaining emergent patterns and spatial heterogeneity. To implement ABM, more data on socio-demographic characteristics and consumer preferences is needed to simulate individual decision-making. Likewise, with the zonal implementation of the SD model, applying the SD model in an ABM can be done relatively easily. However, computation times could increase significantly, as the SD process must take place more than a hundred thousand times per iteration instead of once.
- All preference calculations are based on the results of the last iteration. It can be speculated that consumers will have a recollection of multiple deliveries. Implementing a (moving) model memory can account for this aspect.
- All three scenarios start with two to four iterations where the preference distribution fluctuates to a stable situation. To make the model behaviour consistent over the entire simulation time, a start-up period can be introduced. This behaviour also indicates that the initial estimate preference

distribution is not in-line with the estimates model. The model expects a large share of consumers to choose self-collection than stated preference and revealed preference research shows. A reason, therefore, could be that the weight of the pick-up distance attribute is too low or consumers regard additional aspects of self-collection that are not taken into account in this study.

- The Bass diffusion model is applied to predict the WoM effect. However, that method is designed for a scenario that does not totally reflect the system of this study. The main assumption is that all potential adopters will purchase that service. In this system, a new method could be introduced, or the excellent performance of one service could be below average in later time points. In that changing environment, that assumption is not valid. Secondly, Bass's theory is intended for new innovations. However, van and self-collection delivery are already well-known services. Therefore, it could be the case that the WoM formulation differs for those delivery methods and crowdshipping & drones.
- The impact of WoM is quite limited, a result of linking the Bass diffusion model with the relative performance of the delivery methods. For example, in the formulation of Bass (2004), yearly sales growth rates can reach 70% to 90% in the first years. And even though that formulation incorporates the effect of innovators, the WoM rates in this study are just a few percentages. More accurate knowledge about the magnitude of the WoM effect could increase the model prediction quality.
- The magnitude of familiarity is assumed to be comparable with the WoM effect. Also, more accurate magnitude estimations could increase the model prediction quality in this case. Thereby, the magnitude of familiarity is based on the system's memory over all iterations. Again, a moving memory could be more justifiable as a five-year period is relatively long, certainly in a changing environment.
- The self-collection point assignment is only based on consumer demand. Considerations from the carrier perspective are not regarded, and their decision rule could lead to different results. Additionally, the model formulation for the self-collection evolution only considers the distance score. The results show that the distance score indeed reduces slightly. However, the reliability score increased substantially, thereby worsening the performance score of self-collection significantly. A model that considers both attributes could potentially lead to a better compromise. Furthermore, many points are removed during the iteration period. While in reality, the investment in implementing a self-collection point is quite high and therefore, an already established self-collection point will not be removed swiftly. Lastly, self-collection points are placed within the centre of a zone. Placing those points on the edge of two or more zones could improve self-collection performance with fewer self-collection points.
- The potential of crowdshipping does depend on the willingness to crowdship, which is kept constant in this simulation. That willingness is likely to also be dynamic, for example, because of changing remuneration in peak periods or WoM about acting as a crowdshipper.
- The developed drone delivery model is very simple. It does not consider flight regulations or no-fly zones. Furthermore, it is assumed that parcels can be delivered by drone, which is not valid in the real world with heavy or odd-sized parcels.
- A crucial assumption is that the supply of the delivery method is demand based in this model. Carriers could apply other decision rules, for example, because of competitive strategies or personnel limitations. Future research could explore the strategic decision-making of multiple carriers and the implications of such an approach.

6.2.3. Future Research

Aside from model improvements, there are various recommendations to improve the understanding of consumer preferences for last-mile delivery solutions.

- Available data on consumer preferences for different delivery methods is scarce. The studies that explored this field often use different attributes and delivery methods. Consequently, comparing different studies is complex and general assumptions; for example, attribute beta weight ranges do not exist. Additionally, limited data is available on the delivery method that consumers opt for and why they do this. The MPN data on which the MASS-GT parcel estimation is based does not collect data about preferences and expectations. It is recommended to add these questions during a new data collection effort.
- In many studies, parcels are assumed to be delivered within a specified timeframe. However, failed deliveries occur frequently, are costly for a parcel carrier and disrupt the consumer experience. In this study, the model was shown to be noticeably sensitive to the reliability attribute as it has a strong dynamic relation between consumer preferences and operational performance. Therefore, it is suggested to include reliability in further research.
- The performance score is closely related to a utility function. It is recommended to perform an SP study with the delivery methods and attributes chosen in this study. With that data, an actual utility function can be estimated that provides more accurate forecasting. Additionally, the attribute levels in this study are based on usage in literature. However, in, for example, the Dutch context, current delivery is generally offered for free, and the most expensive delivery result in this study is below €2.50. Smaller levels for the costs attribute could be more insightful in this case.
- The developed model can be used to perform extensive scenario analysis. Future research could, for example, explore different governmental regulations.
- In this study, the SD model is used in the case of South Holland. This is a highly urbanised area without 'real' rural zones and an ample supply of depots from diverse carriers. It is thus questionable if results from this region apply in another context. Travel costs and delivery speed could vary strongly in other environments.
- The preferences of consumers are strongly based on the performance score of the delivery methods. That score is grounded on four performance-related attributes. Yet, in literature, for example, the paper of Caspersen and Navrud (2021), it is indicated that consumers take environmental impact into consideration in their decision process. In this model, such an attribute could be implemented in the performance score as well as the WoM effect.
- It is assumed that all parcel requests can be fulfilled from each depot. Further research could assess if, for products that are not generally available at depots and thus have a longer lead time, consumers value the last-mile differently.

6.2.4. Policymakers

Although this study is exploratory for a new methodology to describe consumer preferences for last-mile delivery methods, and thus not an exact forecasting model, the results show various useful insights for policymakers. Firstly, last-mile delivery in South Holland is very dependent on van delivery. Even in the scenario that estimates the lowest market share for van delivery, still, 39% of the consumers would prefer van delivery. Thereby, vans form a backup for parcels that cannot be delivered via crowdshipping of drones. If policymakers or parcel carriers want to reduce the demand for van delivery, providing additional delivery methods could lessen that demand. However, those delivery methods should establish very high service levels, which is not likely or very costly in many zones. A large impact could also be organized by reducing the service of van delivery, mainly by introducing a delivery fee. Currently, van delivery is offered for free in the Netherlands, and because consumers are sensitive to cost increases, a demand shift can be implemented by forcing a price to van delivery.

Policymakers should consider that the main share of van kilometres is made by zone-to-zone travel. Therefore, to significantly reduce these distances, a policy could focus on phasing out van delivery in entire zones. This could be done by requiring parcel carriers to only perform other delivery methods in specific zones.

Furthermore, it is concluded that self-collection is a highly competitive delivery method in dense urban areas. To make it even more successful, municipalities should direct or build a vast network, especially white-label, automated lockers. Those points can provide a convenient performance to consumers while reducing the amount of trips carriers need to perform. Thereby, it provides the policymakers with the opportunity of placing the lockers in such a way that congestion forming can be circumvented.

Both crowdshipping and drone delivery are implemented in this study without constraints from regulations. Thus, theoretically, these new delivery methods can play an important role in the last-mile in South Holland, certainly drone delivery. However, regulations could limit the level of service and the zones where these delivery methods can be operated. Clear regulations, or at least indications of future regulations, could help carriers in their decision-making for the development of new delivery methods.

Bibliography

ACM. (2018). Post- en pakkettenmonitor 2018 [Rapport ACM/INT/390670]. *De Autoriteit Consument en Markt*.

ACM. (2020). Marktstudie last mile pakketbezorging [Rapport ACM/UIT/534379]. *De Autoriteit Consument en Markt*.

ACM. (2021). Post- en pakkettenmonitor 2021 [Rapport ACM/UIT/579837]. *De Autoriteit Consument en Markt*.

ACM. (2022). Marktstudie last mile pakketbezorging. *De Autoriteit Consument en Markt*.

Allen, J., Piecyk, M., Piotrowska, M., McLeod, F., Cherrett, T., Ghali, K., Nguyen, T., Bektas, T., Bates, O., Friday, A., et al. (2018). Understanding the impact of e-commerce on last-mile light goods vehicle activity in urban areas: The case of london. *Transportation Research Part D: Transport and Environment*, *61*, 325–338.

Allsop, D. T., Bassett, B. R., & Hoskins, J. A. (2007). Word-of-mouth research: Principles and applications. *Journal of advertising research*, *47*(4), 398–411.

Amazon hub. (2023). Amazon. https://techcrunch.com/2017/07/27/amazon-launches-the-hub-parcel-delivery-lockers-for-apartment-buildings/

Angerhofer, B. J., & Angelides, M. C. (2000). System dynamics modelling in supply chain management: Research review. *2000 Winter Simulation Conference Proceedings (Cat. No. 00CH37165)*, *1*, 342–351.

Asdecker, B. (2020). How dare you replace my deliveryman?! acceptance of last-mile transportation innovations–a qualitative perspective.

Aurambout, J.-P., Gkoumas, K., & Ciuffo, B. (2019). Last mile delivery by drones: An estimation of viable market potential and access to citizens across european cities. *European Transport Research Review*, *11*(1), 1–21.

Average co2 emissions from new passenger cars, by eu country. (2022). ACEA. https://www.acea.auto/figure/average-co2-emissions-from-new-passenger-cars-by-eu-country/

Average co2 emissions from newly sold passenger cars in the netherlands between 2008 to 2019. (2023). statista. https://www.statista.com/statistics/762801/average-co2-emissions-from-new-cars-in-the-netherlands/

Bass, F. M. (2004). A new product growth for model consumer durables. *Management science*, *50*(12 supplement), 1825–1832.

The bass diffusion model. (2023). https://www.ashokcharan.com/Marketing-Analytics/~pv-bass-diffusion-model.php#gsc.tab=0

Beeldbank. (n.d.). *deBuren*. https://www.deburen.nl/beeldbank/

Berendschot, M. (2021). The effects of crowdshipping on transport use.

Bilik, J. (2014). *Parcel machines - green solution for green cities*. Presentation at 1st International Conference Green Logistics for Greener Cities, Szczecin, Poland.

Boysen, N., Fedtke, S., & Schwerdfeger, S. (2021). Last-mile delivery concepts: A survey from an operational research perspective. *Or Spectrum*, *43*, 1–58.

Buijs, P., & Niemeijer, R. (2022). A greener last mile: Reviewing the carbon emission impact of pickup points in last-mile parcel delivery. *Available at SSRN 4169737*.

Buldeo Rai, H., Verlinde, S., & Macharis, C. (2019). The “next day, free delivery” myth unravelled: Possibilities for sustainable last mile transport in an omnichannel environment. *International Journal of Retail & Distribution Management*, *47*(1), 39–54.

Buldeo Rai, H., Verlinde, S., & Macharis, C. (2021). Who is interested in a crowdsourced last mile? a segmentation of attitudinal profiles. *Travel Behaviour and Society*, *22*, 22–31.

Carson, J. S. (2002). Model verification and validation. *Proceedings of the winter simulation conference*, *1*, 52–58.

Caspersen, E., & Navrud, S. (2021). The sharing economy and consumer preferences for environmentally sustainable last mile deliveries. *Transportation Research Part D: Transport and Environment*, *95*, 102863.

Cauwelier, K., Macharis, C., & Mommens, K. M. (2023). Travel behavior of e-consumers: Do travel habits vary among last-mile practices? *Vervoerslogistiek werkdagen 2023*, 515–530.

CBS. (2022). Nederlander koopt meer digitale producten maar minder goederen online. https://www.cbs.nl/nl-nl/nieuws/2022/42/nederlander-koopt-meer-digitale-producten-maar-minder-goederen-online

Chen, C., Leon, S., & Ractham, P. (2022). Will customers adopt last-mile drone delivery services? an analysis of drone delivery in the emerging market economy. *Cogent Business & Management*, *9*(1), 2074340.

Convenant zes. (2020). Gemeente Rotterdam. https://logistiek010.nl/app/uploads/2021/03/Convenant_ZES.pdf

Coolblue start eigen bezorgservice per fiets in belgië: Coolbluefietst. (2018). *Coolblue*. https://aboutcoolblue.com/news/nl-be/coolblue-start-eigen-bezorgservice-per-fiets-in-belgie-coolbluefietst/

D'Andrea, R. (2014). Guest editorial can drones deliver? *IEEE Transactions on Automation Science and Engineering*, *11*(3), 647–648.

de Bok, M., Tavasszy, L., Kourounioti, I., Thoen, S., Eggers, L., Nielsen, V., & Streng, J. (2021). Application of the harmony tactical freight simulator to a case study for zero emission zones in rotterdam [Transportation Research Board Annual Meeting ; Conference date: 05-01-2021 Through 29-01-2021]. *Transportation Research Board Annual Meeting*. http://www.trb.org/AnnualMeeting/AnnualMeeting.aspx

De La Torre, G., Gruchmann, T., Kamath, V., Melkonyan, A., & Krumme, K. (2019). A system dynamics-based simulation model to analyze consumers' behavior based on participatory systems mapping–a "last mile" perspective. *Innovative Logistics Services and Sustainable Lifestyles: Interdependencies, Transformation Strategies and Decision Making*, 165–194.

de Almeide Correia, G. H. (2022). System dynamics approach for modeling urban system complexity [Lecture slides from the course CIE5816 Urban Regions, Transport and Economics].

de Bok, M., Tavasszy, L., & Thoen, S. (2022). Application of an empirical multi-agent model for urban goods transport to analyze impacts of zero emission zones in the netherlands. *Transport Policy*, *124*, 119–127.

Deloison, T., Hannon, E., Huber, A., Heid, B., Klink, C., Sahay, R., & Wolff, C. (2020). The future of the last-mile ecosystem: Transition roadmaps for public- and private-sector players. *World Economic Forum*.

Demir, E., Syntetos, A., & van Woensel, T. (2022). Last mile logistics: Research trends and needs. *IMA Journal of Management Mathematics*, *33*(4), 549–561.

de Oliveira, L. K., Morganti, E., Dablanc, L., & de Oliveira, R. L. M. (2017). Analysis of the potential demand of automated delivery stations for e-commerce deliveries in belo horizonte, brazil. *Research in Transportation Economics*, *65*, 34–43.

Fandos Herrera, C., & Flavián Blanco, C. (2011). Consequences of consumer trust in pdo food products: The role of familiarity. *Journal of Product & Brand Management*, *20*(4), 282–296.

Gatta, V., Marcucci, E., Nigro, M., Patella, S. M., & Serafini, S. (2018). Public transport-based crowdshipping for sustainable city logistics: Assessing economic and environmental impacts. *Sustainability*, *11*(1), 145.

Gatta, V., Marcucci, E., Nigro, M., & Serafini, S. (2019). Sustainable urban freight transport adopting public transport-based crowdshipping for b2c deliveries. *European Transport Research Review*, *11*(1), 1–14.

Grawe, S. J. (2009). Logistics innovation: A literature-based conceptual framework. *The International Journal of Logistics Management*, *20*(3), 360–377.

Hepp, S. B. (2018). *Innovation in last mile delivery: Meeting evolving customer demands: The case of in-car delivery* (Doctoral dissertation).

Ignat, B., & Chankov, S. (2020). Do e-commerce customers change their preferred last-mile delivery based on its sustainability impact? *The International Journal of Logistics Management*, *31*(3), 521–548.

ITF. (2022). The freight space race: Curbing the impact of freight deliveries in cities. *International Transport Forum Policy Papers*, *109*.

Joerss, M., Shcröder, J., Neuhaus, F., Klink, C., & Mann, F. (2016). Parcel delivery: The future of last mile. *McKinsey&Company*.

Johnson, E. J., & Russo, J. E. (1984). Product familiarity and learning new information. *Journal of consumer research*, *11*(1), 542–550.

Keeney, T. (2015). *How can amazon charge $1 for drone delivery?* Ark Invest.

Kim, S. H. (2020). Choice model based analysis of consumer preference for drone delivery service. *Journal of Air Transport Management*, *84*, 101785.

Kox, M. (2021). Drukte bij postnl onverminderd groot, pakketbezorgers trekken aan de bel. *RTLnieuws*. https://www.rtlnieuws.nl/economie/bedrijven/artikel/5207052/postnl-pakketjes-pakketten-post-bezorging-zelfstandigen-beloning

Leon, S., Chen, C., & Ratcliffe, A. (2023). Consumers' perceptions of last mile drone delivery. *International Journal of Logistics Research and Applications*, *26*(3), 345–364.

Mahajan, V., Muller, E., & Bass, F. M. (1990). New product diffusion models in marketing: A review and directions for research. *Journal of marketing*, *54*(1), 1–26.

Mahajan, V., Muller, E., & Bass, F. M. (1995). Diffusion of new products: Empirical generalizations and managerial uses. *Marketing science*, *14*(3_supplement), G79–G88.

Maltese, I., Le Pira, M., Marcucci, E., Gatta, V., & Evangelinos, C. (2021). Grocery or@ grocery: A stated preference investigation in rome and milan. *Research in Transportation Economics*, *87*, 101096.

Mancini, S., & Gansterer, M. (2021). Vehicle routing with private and shared delivery locations. *Computers & Operations Research*, *133*, 105361.

Marcucci, E., Le Pira, M., Carrocci, C. S., Gatta, V., & Pieralice, E. (2017). Connected shared mobility for passengers and freight: Investigating the potential of crowdshipping in urban areas. *2017 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS)*, 839–843.

Martin, R., & Schlüter, M. (2015). Combining system dynamics and agent-based modeling to analyze social-ecological interactions—an example from modeling restoration of a shallow lake. *Frontiers in Environmental Science*, *3*, 66.

Mass-gt. (2023). *Delft University of Technology*. https://www.tudelft.nl/transport/onderzoeksthemas/goederenvervoer-logistiek/sleutelprojecten/mass-gt

Massiani, J., & Gohs, A. (2015). The choice of bass model coefficients to forecast diffusion for innovative products: An empirical investigation for new automotive technologies. *Research in Transportation Economics*, *50*, 17–28.

McCabe, S., Kwan, H., & Roorda, M. J. (2006). Freight transportation: Who is the decision maker. *53rd Annual North American Meetings of the Regional Science Association International*.

Melkonyan, A., Gruchmann, T., Lohmar, F., Kamath, V., & Spinler, S. (2020). Sustainability assessment of last-mile logistics and distribution strategies: The case of local food networks. *International Journal of Production Economics*, *228*, 107746.

Merkert, R., Bliemer, M. C., & Fayyaz, M. (2022). Consumer preferences for innovative and traditional last-mile parcel delivery. *International Journal of Physical Distribution & Logistics Management*.

Mohammad, W. A., Nazih Diab, Y., Elomri, A., & Triki, C. (2023). Innovative solutions in last mile delivery: Concepts, practices, challenges, and future directions. *Supply Chain Forum: An International Journal*, 1–19.

Molin, E., Kosicki, M., & van Duin, R. (2022). Consumer preferences for parcel delivery methods: The potential of parcel locker use in the netherlands. *Eur. J. Transp. Infrastruct. Res*, *22*, 183–200.

Morganti, E., Seidel, S., Blanquart, C., Dablanc, L., & Lenz, B. (2014). The impact of e-commerce on final deliveries: Alternative parcel delivery services in france and germany. *Transportation Research Procedia*, *4*, 178–190.

Nguyen, D. H., De Leeuw, S., Dullaert, W., & Foubert, B. P. (2019). What is the right delivery option for you? consumer preferences for delivery attributes in online retailing. *Journal of Business Logistics*, *40*(4), 299–321.

Ogden, K. (1992). *Urban goods movement: A guide to policy and planning*. Ashgate Publishing Company.

Pani, A., Mishra, S., Golias, M., & Figliozzi, M. (2020). Evaluating public acceptance of autonomous delivery robots during covid-19 pandemic. *Transportation research part D: transport and environment*, *89*, 102600.

Pruyt, E. (2013). Small system dynamics models for big issues: Triple jump towards real-world complexity.

Punakivi, M., Yrjölä, H., & Holmström, J. (2001). Solving the last mile issue: Reception box or delivery box? *International Journal of Physical Distribution & Logistics Management*, *31*(6), 427–439.

Punel, A., Ermagun, A., & Stathopoulos, A. (2018). Studying determinants of crowd-shipping use. *Travel Behaviour and Society*, *12*, 30–40.

Punte, S., Tavasszy, L., Baeyens, A., & Liesa, F. (2019). A framework and process for the development of a roadmap towards zero emissions logistics 2050. *Alliance for Logistics Innovation through Collaboration in Europe*.

Rabe, M., Chicaiza-Vaca, J., & Gonzalez-Feliu, J. (2020). Concept for simulation-optimization procedure model for automated parcel lockers as a last-mile delivery scheme: A case study in the city of dortmund. *13th International Conference of Research in Logistics and Supply Chain Management. July 4th-7th, Le Havre, France*, 765–774.

Rai, H. B., Verlinde, S., & Macharis, C. (2021). Unlocking the failed delivery problem? opportunities and challenges for smart locks from a consumer perspective. *Research in Transportation Economics*, *87*, 100753.

Randall, C. (2019). China: Dhl launches drone delivery service. https://www.electrive.com/2019/05/20/china-dhl-launches-drone-delivery-service/

Ranieri, L., Digiesi, S., Silvestri, B., & Roccotelli, M. (2018). A review of last mile logistics innovations in an externalities cost reduction vision. *Sustainability*, *10*(3), 782.

Revealing the secret emissions of e-commerce; hint: It's all in the delivery. (2022). Clean Mobility Collective. https://clean-mobility.org/wp-content/uploads/2022/07/Secret-Emissions-of-E-Commerce.pdf

Roorda, M. J., Cavalcante, R., McCabe, S., & Kwan, H. (2010). A conceptual framework for agent-based modelling of logistics services. *Transportation Research Part E: Logistics and Transportation Review*, *46*(1), 18–31.

Rózycki, M., Gral, M., Anson, J., & O'Neill, J. (2021). Out-of-home delivery in europe 2021 pudo and parcel lockers.

Rózycki, M., Gral, M., Pit, L., Streule, I., Anson, J., O'Neill, J., & Rai, H. B. (2022). Green last mile europe report 2022.

Sargent, R. G. (2010). Verification and validation of simulation models. *Proceedings of the 2010 winter simulation conference*, 166–183.

Schröder, J., Heid, B., Neuhaus, F., Kässer, M., Klink, C., & Tatomier, S. (2018). *Technology delivered: Implications for cost, customers, and competition in the last-mile ecosystem*. McKinsey&Company.

Seghezzi, A., Siragusa, C., & Mangiaracina, R. (2022). Parcel lockers vs. home delivery: A model to compare last-mile delivery cost in urban and rural areas. *International Journal of Physical Distribution & Logistics Management*.

Self-driving vehicles in logistics. a dhl perspective on implications and use cases for the logistics industry [DHL Trend Research]. (2014).

Shabbir, A., Malik, S. A., & Janjua, S. Y. (2017). Equating the expected and perceived service quality: A comparison between public and private healthcare service providers. *International Journal of Quality & Reliability Management*.

Shepherd, S. (2014). A review of system dynamics models applied in transportation. *Transportmetrica B: Transport Dynamics*, *2*(2), 83–105.

Starship technologies, aramark launch contactless robot food-delivery service at asu. (2020). *Arizona State University*. https://news.asu.edu/20200904-asu-news-starship-aramark-launch-contactless-robot-food-delivery

Sterman, J. (2002). System dynamics: Systems thinking and modeling for a complex world.

Sudbury, A. W., & Hutchinson, E. B. (2016). A cost analysis of amazon prime air (drone delivery). *Journal for Economic Educators*, *16*(1), 1–12.

Tapia, R. J., Kourounioti, I., Thoen, S., de Bok, M., & Tavasszy, L. (2023). A disaggregate model of passenger-freight matching in crowdshipping services. *Transportation Research Part A: Policy and Practice*, *169*, 103587.

Thaller, C., Clausen, U., & Kampmann, R. (2016). System dynamics based, microscopic freight transport simulation for urban areas: A methodological approach. *Commercial Transport: Proceedings of the 2nd Interdiciplinary Conference on Production Logistics and Traffic 2015*, 55–72.

Thoen, S., de Bok, M., Eggers, L., & Kourounioti, I. (2021). E-commerce in strategic freight models: An implementation in the province of zuid-holland [Slides for European Transport Conference 2021].

Thoen, S., de Bok, M., & Tavasszy, L. (2020). Shipment-based urban freight emission calculation. *2020 Forum on Integrated and Sustainable Transportation Systems (FISTS)*, 372–377.

Thoen, S., Tavasszy, L., de Bok, M., Correia, G., & van Duin, R. (2020). Descriptive modeling of freight tour formation: A shipment-based approach. *Transportation Research Part E: Logistics and Transportation Review*, *140*, 101989.

Vakulenko, Y., Hellström, D., & Hjort, K. (2018). What's in the parcel locker? exploring customer value in e-commerce last mile delivery. *journal of Business Research*, *88*, 421–427.

Vakulenko, Y., Shams, P., Hellström, D., & Hjort, K. (2019). Service innovation in e-commerce last mile delivery: Mapping the e-customer journey. *Journal of Business Research*, *101*, 461–468.

van Duin, J. R., van Kolck, A., Anand, N., Taniguchi, E., et al. (2012). Towards an agent-based modelling approach for the evaluation of dynamic usage of urban distribution centres. *Procedia-Social and Behavioral Sciences*, *39*, 333–348.

van Duin, J. R., Wiegmans, B. W., van Arem, B., & van Amstel, Y. (2020). From home delivery to parcel lockers: A case study in amsterdam. *Transportation Research Procedia*, *46*, 37–44.

Verkeersmodel geeft inzicht in verkeersstromen metropoolregio. (2023). https://mrdh.nl/project/verkeersmodel

Walker, J., & Ben-Akiva, M. (2002). Generalized random utility model. *Mathematical social sciences*, *43*(3), 303–343.

Wang, X., Yuen, K. F., Wong, Y. D., & Teo, C.-C. (2020). E-consumer adoption of innovative last-mile logistics services: A comparison of behavioural models. *Total Quality Management & Business Excellence*, *31*(11-12), 1381–1407.

Weltevreden, J. W. (2008). B2c e-commerce logistics: The rise of collection-and-delivery points in the netherlands. *International journal of retail & distribution management*, *36*(8), 638–660.

Wong, R., & Sheng, S. Y. (2012). A business application of the system dynamics approach: Word-of-mouth and its effect in an online environment. *Technology Innovation Management Review*, *2*(6).

A

Methodology Consumer Delivery Preference Literature Review

In this appendix, the methodology for the selection of papers for the research on consumer attributes for parcel delivery is described. For the collection of papers, the database Scopus was used to search for studies about the preferences of customers for the delivery of parcels to their homes. A specific start date was not used, but the search was limited to English and Dutch studies. The used search strategy can be seen in Table A.1. The Scopus database was searched within the article titles, abstracts and keywords. As the selection was done on 28-01-2023 studies published after this data are excluded.

Table A.1: Search strategy and hits on Scopus on 28-01-2023

Step	Truncation	Hits
1	E-commerce AND delivery	2344
2	Last-mile AND delivery	1717
3	E-commerce AND delivery AND preference	135
4	Last-mile AND delivery AND customer AND preference	38
5	Last-mile AND delivery AND consumer AND preference	39
6	Last-mile AND delivery AND (customer OR consumer) AND preference	63

The first two search strings generated quite a large number of hits. As this research tries to asses how customers perceive different delivery methods the search is further limited to preferences. This leads to a more reasonable amount of hits in Scopus. Searching for last-mile delivery narrows the selection of studies even more. As customer and consumer both described people ordering online and they lead to other studies, it is chosen to combine both words in the search string. To again limit the list of literature only the twenty most cited papers on 28-01-2023 are considered. Within this collection papers were selected on title, abstract, and finally the full text. Table A.2 presents an overview of the eight selected papers.

Table A.2: Overview of the selected literature

Reference	Title
Gatta et al., 2018	Public Transport-Based Crowdshipping for Sustainable City Logistics: Assessing Economic and Environmental Impacts
Nguyen et al., 2019	What Is the Right Delivery Option for You? Consumer Preferences for Delivery Attributes in Online Retailing
de Oliveira et al., 2017	What Is the Right Delivery Option for You? Consumer Preferences for Delivery Attributes in Online Retailing
Gatta et al., 2019	Sustainable urban freight transport adopting public transport-based crowdshipping for B2C deliveries
Buldeo Rai et al., 2019	Sustainable urban freight transport adopting public transport-based crowdshipping for B2C deliveries
Ignat and Chankov, 2020	Do e-commerce customers change their preferred last-mile delivery based on its sustainability impact?
Caspersen and Navrud, 2021	The sharing economy and consumer preferences for environmentally sustainable last mile deliveries
Maltese et al., 2021	Grocery or @grocery: A stated preference investigation in Rome and Milan

B

Method Specific Causal Loop Diagrams

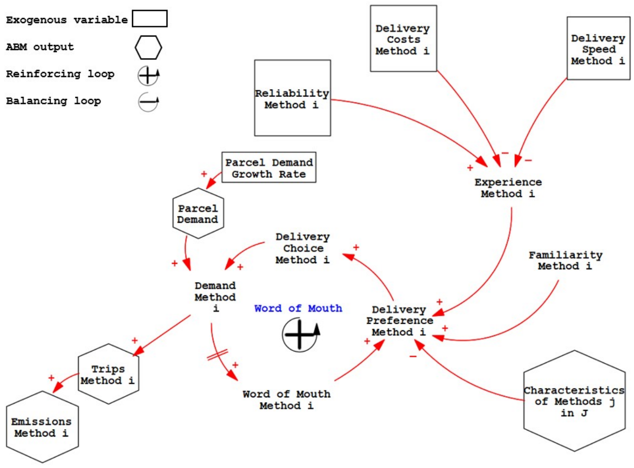


Figure B.1: Causal Loop Diagram for van delivery

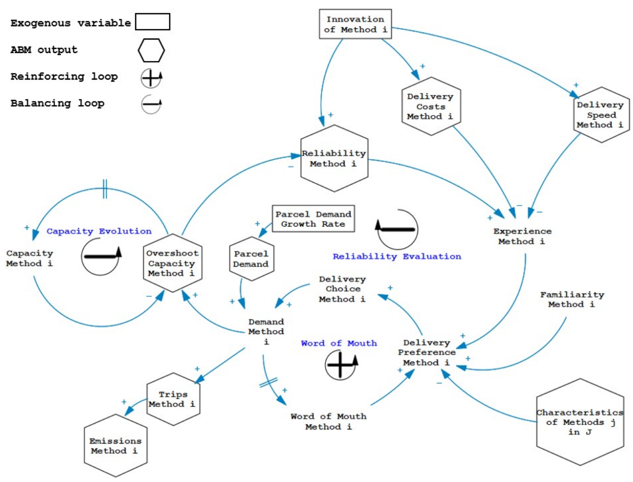


Figure B.2: Causal Loop Diagram for drone delivery

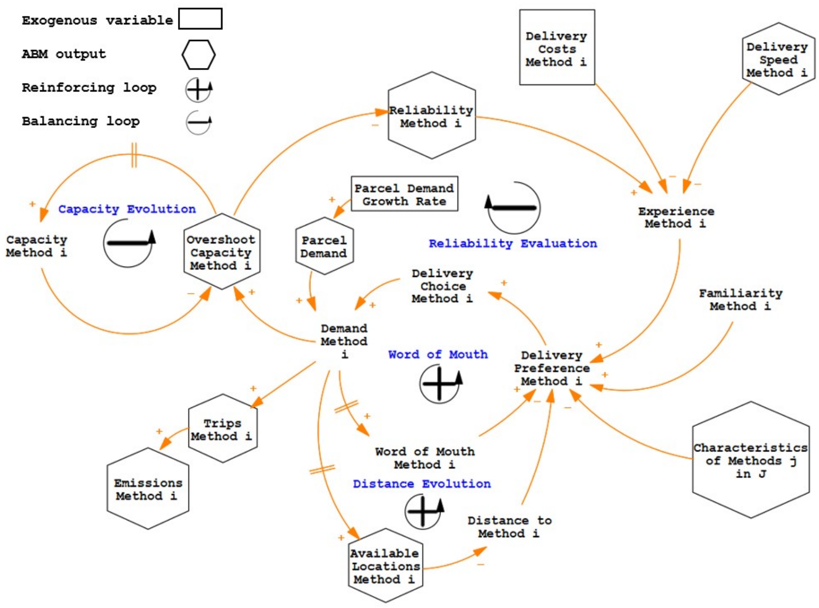


Figure B.3: Causal Loop Diagram for self-collection delivery

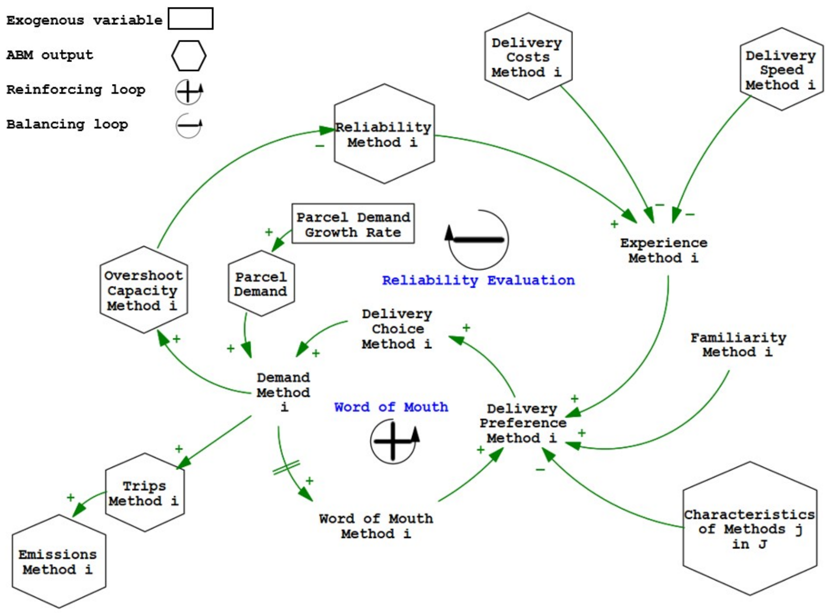


Figure B.4: Causal Loop Diagram for crowdshipping delivery

C

Variables in Simulation Model

This overview shows the variables used in the simulation model. Self-collection is abbreviated to SF and crowdshipping to SW.

Table C.1: Overview of model variables

Variable	Method	Symbol	Unit	Type of Variable	Initial Value	Source
Parcel Demand Growth Rate			%/year	Rate	21.7	ACM, 2017-2021
Parcel Demand			Parcels/day	Stock	115608	Estimation on MPN data
Delivery Preference Method i	Van SF CS Drone		% % % %	Auxiliary Auxiliary Auxiliary Auxiliary	59 20 21 0	See section 4.2 See section 4.2 See section 4.2 See section 4.2
Delivery Choice Method i	Van SF CS Drone		Parcels/day Parcels/day Parcels/day Parcels/day	Rate Rate Rate Rate		N.A. N.A. N.A. N.A.
Demand Method i	Van SF CS Drone		Parcels Parcels Parcels Parcels	Stock Stock Stock Stock		N.A. N.A. N.A. N.A.
Word of Mouth Method i	Van SF CS Drone		% % % %	Auxiliary Auxiliary Auxiliary Auxiliary		N.A. N.A. N.A. N.A.
Available Locations Method i	SF		Location	Stock	1506	See subsubsection 4.5.1.2
Distance to Method i	SF		m	Auxiliary		N.A.

Continued on next page

Table C.1 – continued from previous page

Variable	Method	Symbol	Unit	Type of Variable	Initial Value	Source
Overshoot Capacity Method i	SF		Parcels	Auxiliary		N.A.
	Drone		Parcels	Auxiliary		N.A.
Capacity Method i	SF		Parcels	Stock		N.A.
	Drone		Parcels	Stock		N.A.
Reliability Method i	Van	μR_i	Failed Delivery Rate	Exogenous	25%	Buldeo Rai et al., 2019
	SF	μR_i	Score	Auxiliary	0	N.A.
	CS	μR_i	Score	Auxiliary	0	N.A.
	Drone	μR_i	Score	Auxiliary	0	N.A.
Delivery Costs Method i	Van	μC_i	Score	Exogenous	0	See subsubsection 4.5.4.3
	SF	μC_i	Score	Exogenous	0	See subsubsection 4.5.4.3
	CS	μC_i	Score	Auxiliary	1	See subsubsection 4.5.4.3
	Drone	μC_i	Score	Auxiliary	2	See subsubsection 4.5.4.3
Delivery Speed Method i	Van	μS_i	Score	Exogenous	2	See subsubsection 4.5.4.3
	SF	μS_i	Score	Auxiliary	1	See subsubsection 4.5.4.3
	CS	μS_i	Score	Auxiliary	2	See subsubsection 4.5.4.3
	Drone	μS_i	Score	Auxiliary	0	See subsubsection 4.5.4.3
Experience Method i	Van	Sm_i	Score	Auxiliary		N.A.
	SF	Sm_i	Score	Auxiliary		N.A.
	CS	Sm_i	Score	Auxiliary		N.A.
	Drone	Sm_i	Score	Auxiliary		N.A.
Familiarity Method i	Van		%	Auxiliary		N.A.
	SF		%	Auxiliary		N.A.
	CS		%	Auxiliary		N.A.
	Drone		%	Auxiliary		N.A.
Innovation of Method i	Drone	IP_i	%	Exogenous		N.A.
Trips Method i	Van		Trips	Stock		N.A.
	SF		Trips	Stock		N.A.
	CS		Trips	Stock		N.A.
	Drone		Trips	Stock		N.A.
Emissions Method i	Van		CO2/day	Stock		N.A.
	SF		C02/day	Stock		N.A.
	CS		C02/day	Stock		N.A.

Continued on next page

Table C.1 – continued from previous page

Variable	Method	Symbol	Unit	Type of Variable	Initial Value	Source
Innovation chance	Drone	P_i	%	Exogenous	1	N.A.
Capacity	Van		Parcels/trip	Exogenous	180	Thoen, de Bok, and Tavasszy, 2020
	SF		Parcels/day	Auxiliary	42	*Amazon Hub*, 2023
	Drone		Parcels/trip	Exogenous	1	Boysen et al., 2021
Familiarity constant		Ω_{change}	%	Exogenous	5	N.A.
Vehicle costs	Drone	C_{drone}	Costs/km	Exogenous	0.10	D'Andrea, 2014
Delivery Success Rate	Van		%	Exogenous	75	Buldeo Rai et al., 2019
	SF		%	Exogenous	100	Boysen et al., 2021
	CS		%	Exogenous	75	Buldeo Rai et al., 2019
	Drone		%	Exogenous	90	N.A.
Drop off time	Van		time	Exogenous	120	Thoen, de Bok, and Tavasszy, 2020
	SF		time	Exogenous	300	N.A.
	Drone		time	Auxiliary	120	N.A.
Kilometre CO2 emissions	Van		g/km	Exogenous	213, 126, 185	When full; urban, rural and highway (Thoen, de Bok, & Tavasszy, 2020)
Range	Drone		km	Auxiliary	10	D'Andrea, 2014
Speed	Drone		km/h	Auxiliary	45	D'Andrea, 2014
Profit margin	Drone		%	Exogenous	15	See subsubsection 4.5.4.4
Characteristics of other Methods	Van		%	Auxiliary		N.A.
	SF			Auxiliary		N.A.
	CS			Auxiliary		N.A.
	Drone			Auxiliary		N.A.

End of Table

D

Performance Scores

D.1. Formula Development

The logit formulation as described in subsubsection 4.5.4.3 was not implemented immediately in this study. The performance scores that are calculated for each attribute are not official utilities because this study did not estimate the utility for the different attributes in the study area. And at the initial state of this thesis, it was unclear if sufficient substantiation could be found for the weights of the betas to use the logit formulation reliably. In this subsection two alternative methods are discussed that estimated the probability of preferring a delivery method.

The logit probability relies on the magnitude of the utility of the logit formulation, while initially only the relative performance was known. Therefore, the calculated logit probability is seen as an indication of how a certain delivery method will grow in these alternative formulations. It is assumed that a high logit probability accelerates the growth of the probability of choosing a certain delivery method. This is modelled by calculating for each delivery method the logit probability of choosing that delivery method and multiplying that with the probability of choosing method i in that time point, see Equation D.1. To maintain a summed total preference of 100%, the preference chance for each delivery method is divided by the sum of all preference chances.

$$P_i(t+1) = \frac{P_i(t) * \frac{e^{1-Sm_{it}}}{e^{\sum_{i=0}^{N} 1-Sm_{it}}}}{\sum_{i=0}^{N}\left[P_i(t) * \frac{e^{1-Sm_{it}}}{e^{\sum_{i=0}^{N} 1-Sm_{it}}}\right]} \quad \forall \quad i \in N \tag{D.1}$$

Originally another method was formulated to model the preference evolution. In Equation D.2 the probability is modelled by calculating for each delivery method the relative difference from the mean score and multiplying that with the probability of choosing method i in that time point, see Equation D.2. Which is again summed to a total preference of 100%.

$$P_i(t+1) = \frac{P_i(t) * \left(1 + \frac{\sum_{i=0}^{N} Sm_{it}}{N} - Sm_{it}\right)}{\sum_{i=0}^{N}\left[P_i(t) * \left(1 + \frac{\sum_{i=0}^{N} Sm_{it}}{N} - Sm_{it}\right)\right]} \quad \forall \quad i \in N \tag{D.2}$$

In Figure D.1 it is shown that both formulations perform very comparable. As the logit formulation is a common way of formulating discrete choice probability, the logit model is selected. A disadvantage of this formulation compared to a normal logit model, as in Equation 4.6, is that the model will eventually grow to one delivery method with 100% users and all other delivery methods declining to 0%. Yet, to reach that point a large number of iterations need to take place, this point will not be reached in the application of this study.

These formulations were rejected because of two reasons. The first reason is that a very slight difference in performance will always result in consequent growth for the well-performing delivery method

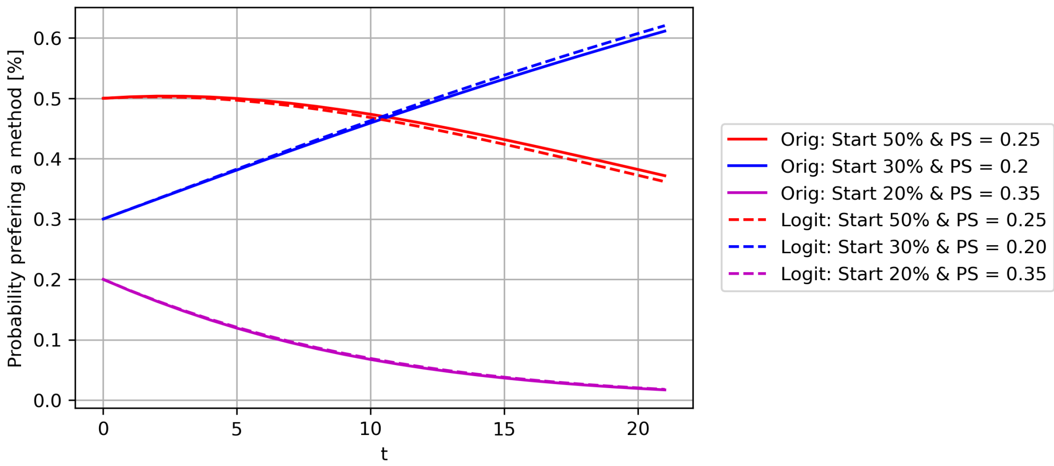


Figure D.1: Influence of the performance score on the probability of preferring a specific method for the original formula and a logit-based formula

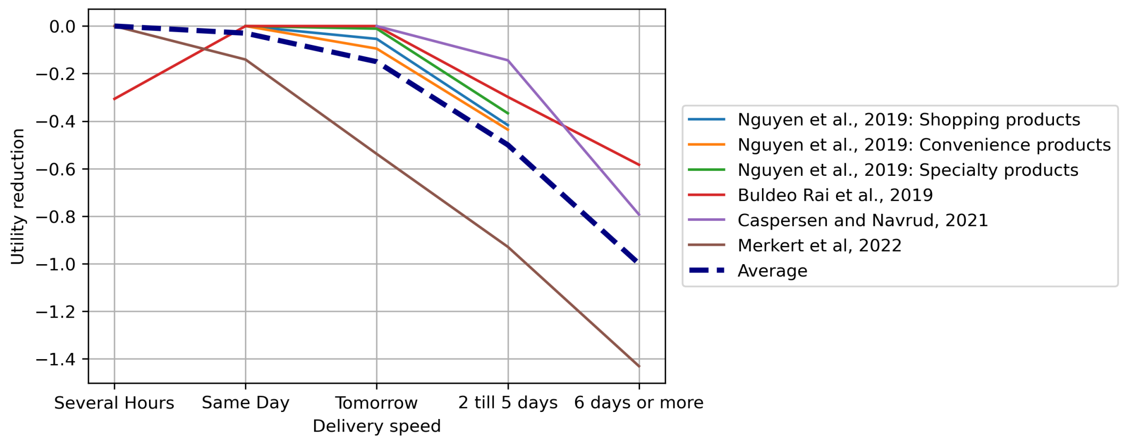
and the opposite for the worse delivery method. Secondly, over time one delivery method will grow to almost 100% probability, while all other delivery methods converge to zero. These characteristics are not in line with other studies. Consequently, it is chosen to use the classical logit formulation, because it is a closer representation of reality. Thereby using the same formulation as other research makes comparing the results more straightforward.

D.2. Level Scores

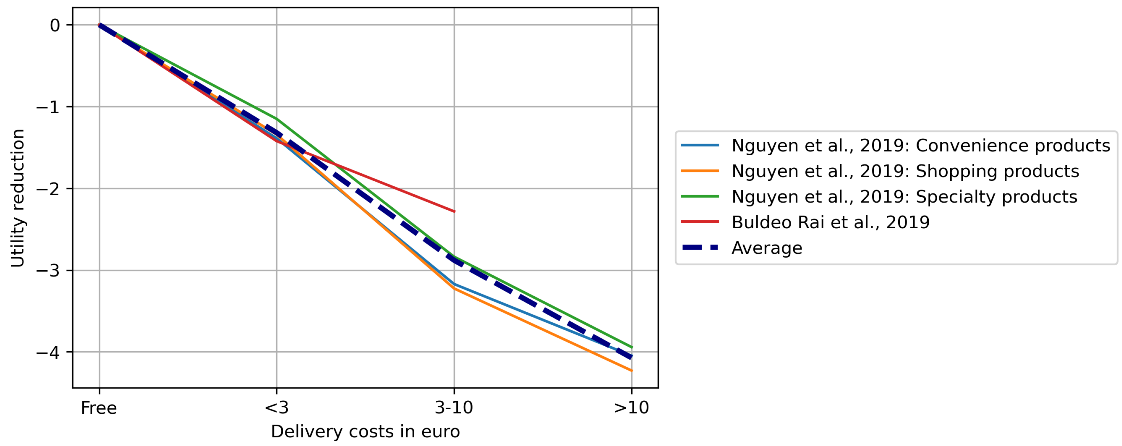
The four attributes in this study consist of multiple levels, these levels have a corresponding performance score that is based on previous research. All attribute scores are normalised, to make comparing the different attributes straightforward. The delivery speed will have exponential growth, which indicates that consumers care less about a delivery tomorrow instead of today than about a delivery taking several days. In Figure D.2a the utility for the delivery speed found in four different studies is plotted and the average line is used as the base for the performance score for each level of the speed attribute.

The cost attribute will have linear distribution, as the average utility is almost linear across the different price ranges in Figure D.2b. Additionally, Molin et al. (2022) concluded that utility decreases linearly in the price range of 0 to 6 euros. As described in chapter 3 reliability of last-mile delivery is not often used as an attribute, it is commonly assumed that a parcel will be delivered to the consumer. Therefore an assumption must be made about the level scores for the reliability attribute. It is believed that a parcel that is not delivered is considerably worse than receiving a parcel with delay or by another delivery method than intended. Consequently, the score of this attribute will be distributed exponentially.

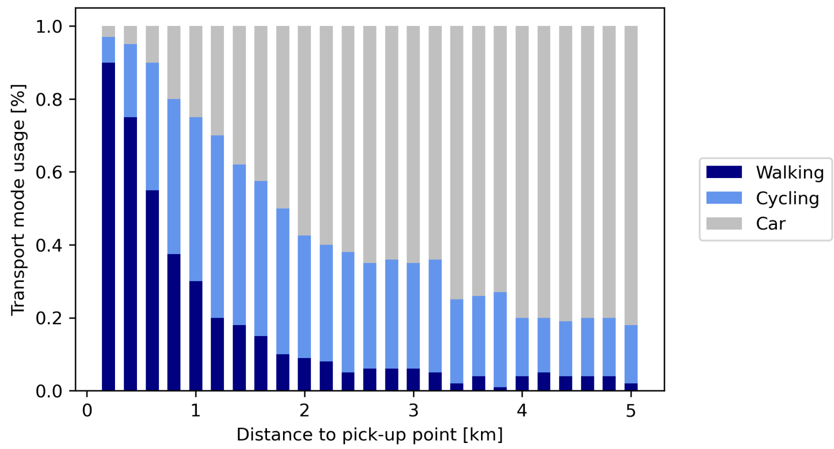
For the distance to a self-collection point, a utility for distinct levels is not provided in earlier work, however, Figure D.2c shows how the use of walking or the bike declines if the distance to a pick-up point grows. It is assumed that the utility of a self-collection point decreases in a similar way as the usage of walking or cycling to pick up a parcel. Therefore the distance attribute will have a logarithmic distribution.



(a) Utility reduction due to the delivery speed in four studies



(b) Utility reduction due to the delivery costs in two studies



(c) Mode of transport per distance to a pick-up point up to 5 kilometres (Buijs & Niemeijer, 2022)

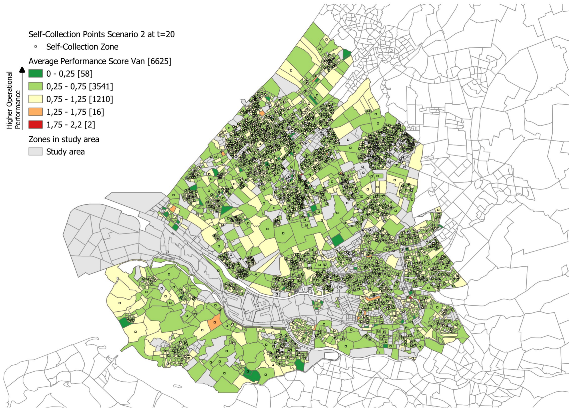
Figure D.2: Sources for the attribute level scores

E

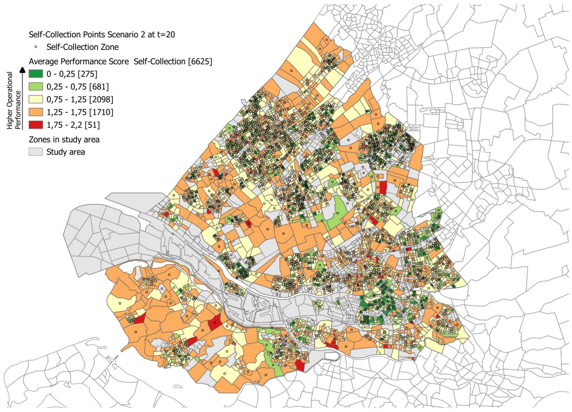
Additional Charts and Graphs

In this appendix, charts and graphs are shown that provide additional insight into the results presented in chapter 5.

Scenario 2



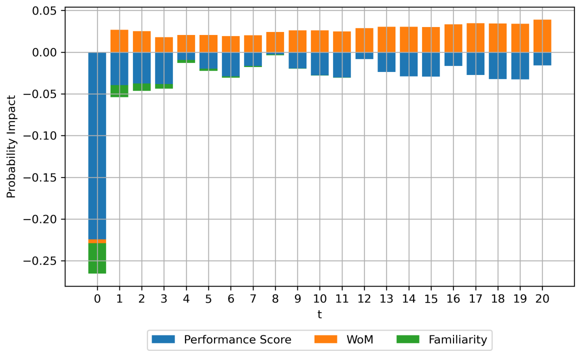
(a) Average performance score Van



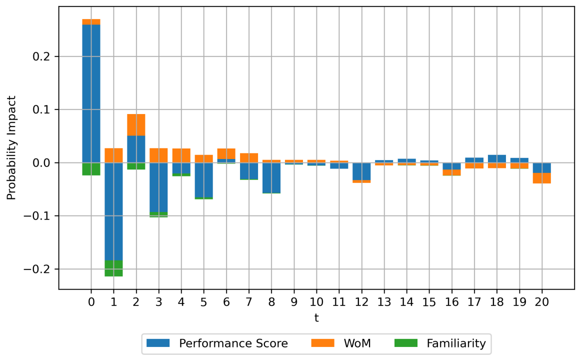
(b) Average performance score Self-Collection

Figure E.1: Average zonal performance score in run 1 scenario 2

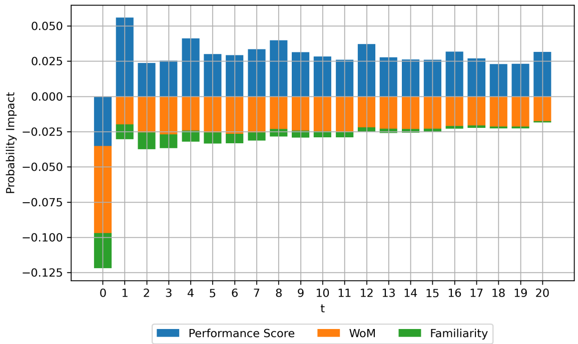
Scenario 3



(a) Impact on the preference probability for Van

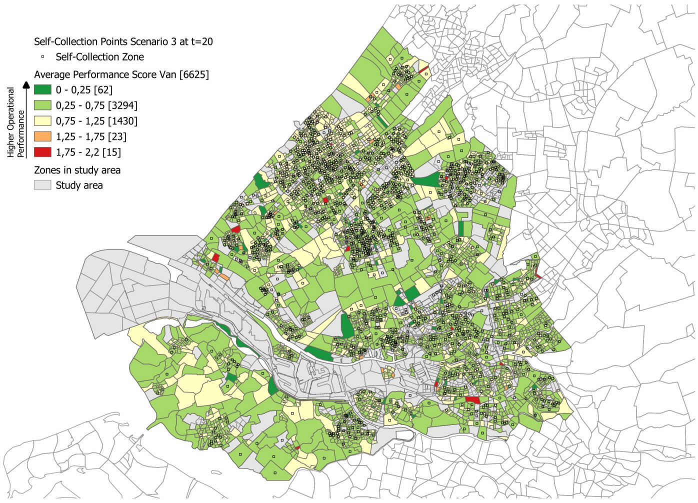


(b) Impact on the preference probability for Self-Collection

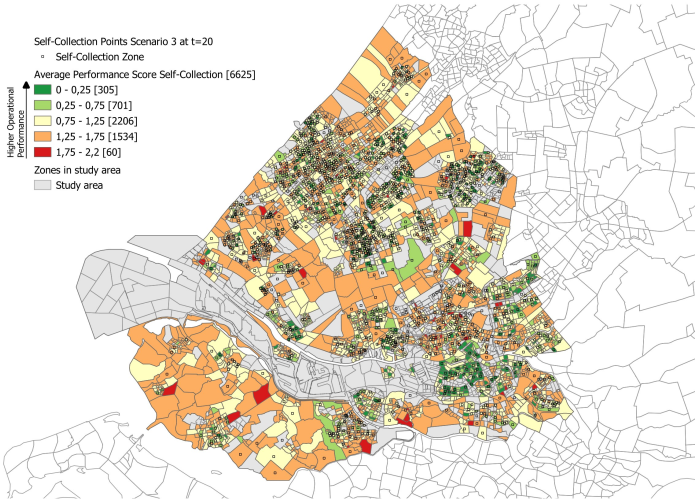


(c) Impact on the preference probability for Crowdshipping

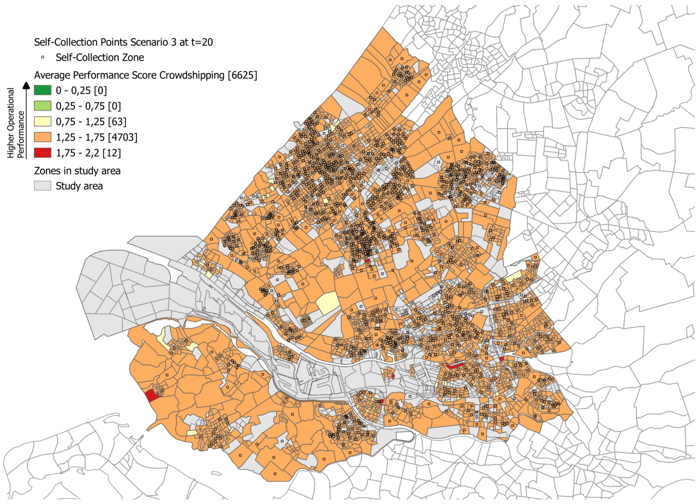
Figure E.2: Preference probability impact in run 1 scenario 3



(a) Average performance score Van



(b) Average performance score Self-Collection



(c) Average performance score Crowdshipping

Figure E.3: Average zonal performance score in run 1 scenario 3

F
Scientific Paper

Evolution of Consumer Preferences in Last-mile Delivery Methods and the Impacts on City Logistic Freight Traffic - A Simulation Study

J. E. van Vliet, G. Homem de Almeida Correia, B. Atasoy and M. A. de Bok
Department of Transport, Infrastructure and Logistics, Delft University of Technology, Delft, The Netherlands

***Abstract*—The growing demand for parcel deliveries is causing significant problems for freight logistics, particularly in cities. The increasing number of delivery vehicles leads to traffic congestion, high emissions, and rising costs. To address these problems, new and sustainable delivery methods must be implemented for the parcel last-mile. However, estimating the impact of a different logistics system is complex, as it depends on consumer adoption of these new delivery methods.**

This paper presents a new simulation model that captures and explores the interconnections between multiple delivery methods and consumer preferences for those delivery methods. Consumers' reaction to the performance and availability of the delivery methods is simulated, next to knowledge progress via word of mouth and familiarisation. The developed hybrid model uses system dynamics to simulate the evolution of consumer preferences and the last-mile delivery via van, self-collection, crowdshipping and drone delivery at an aggregate level, while an agent-based model provides input on the operational performance of the delivery methods. With this structure, consecutive interactions can be simulated, and by that, data on consumer preferences and the delivery method operations can be obtained at multiple time points.

Results of a case study on the province of South Holland show that consumers change their preferences due to the introduction of new delivery methods. However, the vehicle kilometres and emissions of van delivery do not reduce at the same rate as the demand. The study also highlights zonal differences in last-mile services and the importance of delivery quality to consumers.

***Keywords*—Last-mile, consumer preference, freight logistics, e-commerce, system dynamics**

I. Introduction

The Paris Agreement aims to reduce CO2 emissions by at least 50% in 2030 [1]. However, a complex problem arises as a tripling of the global transport demand by 2050 is estimated. A doubling of carbon emissions is expected if the current manner of transportation does not innovate. Last-mile delivery is the part of the supply chain where the consequences, like pollution, congestion and noise, for the freight logistics have the largest impact [2]. Five problems in the last-mile that must be dealt with are [3]: 1) an increasing volume; 2) sustainability; 3) costs; 4) time pressure; 5) an ageing workforce. In addition, DHL marks the last-mile as the 'least predictable part of the entire journey' [4]. The logistics sector seeks to innovate in order to solve the last-mile problem with emerging delivery methods, new delivery locations and automation. Conventional delivery vans are being electrified or can potentially be replaced with other modes [5]. These innovations could provide hard benefits, like reduction of costs, and soft benefits, such as competitive differentiation and service enhancements between parcel carriers [6].

Consumers' acceptance of new and unfamiliar delivery methods is crucial for the success of these delivery innovations [7]. Because of the major role of consumers in the demand and acceptance of delivery methods, it is complex to predict which last-mile innovations provide the most sustainable solution within the parcel freight logistics system [8]. Yet research, especially empirical-based, on consumer preferences is scarce. Likewise, the integration of multiple different delivery methods in one consumer preference model is limited [3]. The problem with the conventional revealed preference (RP) or stated preference (SP) survey data is that they cannot measure future preferences (RP-data) or consider the dynamics between demand levels and capacity and the development of new solutions (SP-data). This creates a gap in how to determine the evolution of consumer preferences while taking into consideration the dynamics in the parcel delivery system.

The objective is to develop a method that simulates to what extent consumer preferences evolve over time due to the performance of conventional (van and self-collection) and emerging (crowdshipping and drone) delivery methods in a spatial model. For that purpose, influential factors for consumer preference development are discussed and modelled, with the added novelty of gathering empirical data on the direction and magnitude of those factors at multiple time points. Furthermore, the evolution of the supply of delivery methods is replicated. This approach allows the exploration of the complex last-mile system where multiple delivery methods compete and complement each other. A case study on the Dutch province of South Holland is performed to generate predictions on real word operations.

This paper is structured as follows. In section II, a literature review is provided on the last-mile, the parcel delivery methods, the theories to describe consumer preferences and system dynamics. The methodology is elaborated on in section III, and the case study is discussed in section IV. The conclusions are presented in section V. Lastly, section VI recommendations on future work are discussed.

II. Literature Review

A. The Last-Mile

The work in [7] defines last-mile delivery as the final transport part of the supply chain from the last distribution centre, consolidation point or warehouse to the consumer. Hence, the last-mile starts after long-haul transportation and ends when the parcel has successfully reached the consumer's preferred destination [3]. The last-mile is further characterised by low-density transport, which is one of the reasons for the high costs, inefficiency and emissions that come with this logistic problem [9]. As the logistics network is reaching its capacity limit, delivery firms must create new delivery methods [10].

B. Delivery Methods

The parcel last-mile to consumers is dominantly performed with human-driven vans [3]. Van delivery has large disadvantages, such as low efficiency, the dependency on the consumer being at home and the reliance on and the cause of traffic. Nevertheless, other delivery methods are not yet disrupting the last-mile environment. Electrification is used to reduce the environmental impact of vans [11]. And cargo bikes are an increasingly popular and environmentally friendly alternative to vans in urban areas. In Europe, pick-up points are a well-known alternative for at-home delivery [12], and the addition of automated lockers boosts the flexibility for consumers [3]. However, only 18% of the parcels in the Netherlands are delivered to a self-collection point [12].

New delivery methods are under development, and real-world test cases are being executed with drones and droids [3]. Both these delivery methods use automation to reduce the operational costs of the carrier. Thereby, the prediction is that drones and droids can offer fast and flexible delivery that is convenient for the consumer [13]. Furthermore, drones open up new possibilities as they are not dependent on road infrastructure. However, both delivery methods still need to overcome technical and especially regulatory barriers. Another disadvantage is the low capacity of these modes. Consequently, in many cases, consolidation centres will be needed [14]. Crowdshipping is another development that can reduce the transport burden of carriers, as it uses the crowd to move or deliver parcels during their own travel [15][16]. A crowdshipper gets remuneration for their work/effort [11], which provides the carrier with a highly flexible and scalable workforce. In this case, drawbacks could be inducing traffic, safety concerns and the dependency on the willingness to act as a crowdshipper [17]. New innovations are also introduced for unattended home delivery, which eliminates the problem of consumers not being at home [3].

C. Evolution of Consumer Preferences

There are various theories that can be used to describe how consumer preferences evolve. With the concept of perceived service quality, the satisfaction of consumers can be assessed [18]. Subsequently, that satisfaction can be used to simulate a delivery preference at a future time point. This theory suggests that consumers confirm their satisfaction by matching their expectations and their experience. The received service consists of soft elements, like reliability and communication, and hard elements, like the quality of the product/service and the proof of performance [6].

Subsequently, the diffusion of innovation (DOI) theory considers a direct relationship between the perceived characteristics of an innovation and the consumer adoption decision [19] [20]. Hence, it can be used to link the perceived service quality to consumer behaviour. DOI stems from marketing, where it is used for the analysis and evaluation of life cycle dynamics [21], and additionally, it is applied for demand forecasting of new products. The DOI theory is often complemented by attitude theories. Instead of a direct relationship, attitude theories suggest that attitude is included as a mediator between beliefs and adoption intention.

The logistics innovation theory describes that the market share of a firm can increase due to a more effective logistics operation [20]. When firms identify a disadvantage with respect to another competing firm due to innovation, those firms will seek to establish the same innovation [19]. In that way, innovations penetrate within that sector. Furthermore, logistics innovations often provide enhanced customer value. The evolution of radical innovations generally shows S-curve behaviour for the consumer benefits of that innovation. Consumer benefits increase slowly at the introduction of an innovation. As the innovation develops, those benefits increase substantially, and finally, the beneficial growth slows as the innovation matures.

Word of mouth (WoM) is an important concept for studies that describe the process of consumers evaluating a service or product and communicating that experience with other consumers [21]. It is reported that for adopting new businesses or services, consumers rely significantly on the experiences and opinions of other consumers in their decision process [7] [22]. WoM influences factors like satisfaction, loyalty, service level and trust. Furthermore, WoM is suitable when a population is heterogeneous or when the interactions between individuals are complex. The Bass diffusion model is a frequently used formula describing that a consumer's initial purchase is related to the number of previous users and that sales of new products grow to a peak and then level to a value lower than that peak [23]. This behaviour results from the growing group of people that can spread WoM due to their experience and a shrinking group of consumers that can still adopt the service. The Bass diffusion model theorises that an innovation is adopted by a group of innovators and a group of imitators. The first group makes their choice independently of others, while imitators base their choice on the social system around them.

Familiarity is defined in [24] as 'the number of product-related experiences that the consumer has accumulated'. It describes the understanding of a product and its characteristics by a consumer and the ability of the consumer to evaluate the quality of a product. A higher level of familiarity with a product tends to make consumers more trustworthy and loyal to that product. Additionally, the introductory experience with a new service is essential for the future perception and expectation of that service [22]. However, consumers also self-educate over time, which can ease initial mistrust or concerns. The familiarity theory can be used to model the resistance or inertia of consumers to choose an unknown delivery method.

D. System Dynamics

System dynamics (SD) links qualitative and quantitative models via a causal loop approach [25] and was developed in 1950-1960 by J. W. Forrester. It is based on the assumption that the behaviour of a system is mostly dependent on its own structure [26]. Domains where SD modelling is used are, for example, health policy, resources scarcity and supply chain management.

The objective of SD is to explore the effect of decisions in a dynamically complex system, where SD supports exploring and simulating non-linear feedback structures and functions [27]. This is accomplished by connecting different system components and linking those connections with mathematical models. The work in [28] identified that the essential viewpoint of SD is that feedback and delay cause the behaviour of systems.

SD describes a system at the aggregate level. It is a much-chosen representation of the internal process of an entity [29], which could be an individual or a group. An advantage of this method is that it can visualise interdependencies between different parts of the system in a transparent way [27]. SD models can be used to perform medium- and long-term forecasts, trend analysis and impact assessments. A benefit of this method is the low data requirement, which is due to the high aggregated level of the model. Furthermore, it can integrate modelling algorithms from multiple fields, like economics and transport modelling. A drawback is that SD does not allow traffic assignment and point-in-time forecasts.

An alternative is agent-based modelling (ABM), which is regularly used to study adaptive systems with self-organisation, emergence and adaptation [29]. ABM helps to explain emergent patterns on a system level while considering the heterogeneity of entities, spatial and temporal heterogeneity, and stochasticity. But these possibilities come with high data requirements to ground a model at the micro-level.

III. Methodology

To be able to simulate the evolution of consumer preferences and the operational fulfilment of the deliveries, a hybrid model is proposed that combines system dynamics (SD) modelling and agent-based modelling (ABM). The SD model will be used to model the preference evolution of consumers over time. The ABM simulates the carriers delivering the parcel from the depot to the receiver and the evolution of performance characteristics of each delivery method. In Fig. 1, the proposed hybrid model is illustrated.

The SD model and the ABM interact in two ways. One interaction is between the delivery operation, which follows from the carriers (agents) performing the parcel delivery, and the consumers evolving their preferences at the system level, resulting in a new demand for the carriers. The second interaction between the SD and AB models specifies that carriers can evolve their operation for each delivery method based on the demand for that delivery method. An advantage of this structure is that consecutive interactions can be simulated. Hence, data on consumer preferences and the delivery method operation can be obtained at multiple time points.

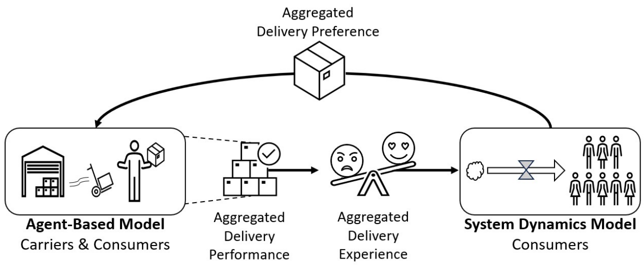


Fig. 1. Schematic overview of the connection between the SD and ABM model

Additionally, the ABM can account for stochasticity in the delivery operations, which creates the possibility to model the consumer response to changes in the upfront chosen delivery.

In this section, the selection of delivery methods and attributes to describe those delivery methods are presented. Subsequently, the SD model is described with a causal loop diagram, see Fig. 2, that simulates the relations between different factors in this consumer-centred, last-mile environment is discussed. Thirdly, the ABM is described.

A. The Delivery Methods

Considering that the focus of this study is on consumer preferences, the delivery methods that provide a distinctive service from the consumer perspective will be regarded. Therefore, this study will simulate van, self-collection, crowdshipping and drone delivery. Self-collection is the combination of pick-up points and automated locker, a simplification that is justifiable as consumers regard both options as quite similar [30]. Crowdshipping provides a unique delivery experience, as parcels are not handled by dedicated carrier personnel. Thereby, the crowdshipping service relies on the crowds' willingness to ship. Both drones and droids offer consumers automated delivery with short delivery times. Drone delivery is selected to explore the potential of a delivery method that is independent of the road infrastructure.

B. Attributes in Consumer Preferences

Key attributes describing consumer parcel delivery choices and/or preferences are identified in [7] [8] [9] [15] [31] [32] [33] [34]. Four attributes are selected: delivery speed, costs, reliability, and pick-up distance. Delivery speed and costs are consistently used in state preference studies. Reliability expresses if the delivery is performed as expected. This attribute is seldom used, as failed deliveries are often not accounted for. Offering consumers a time slot choice is a commonly used attribute that relates to reliability. However, it still disregards the possibility of unsuccessful or disturbed operations. The pick-up distance is only applied to self-collection and accounts for the distance a consumer needs to travel to pick up a parcel.

C. The System Dynamics Model

SD is used to model the evolution of consumer preferences. In Fig. 2, the principal factors and their interactions are displayed in a causal loop diagram. As the proposed simulation

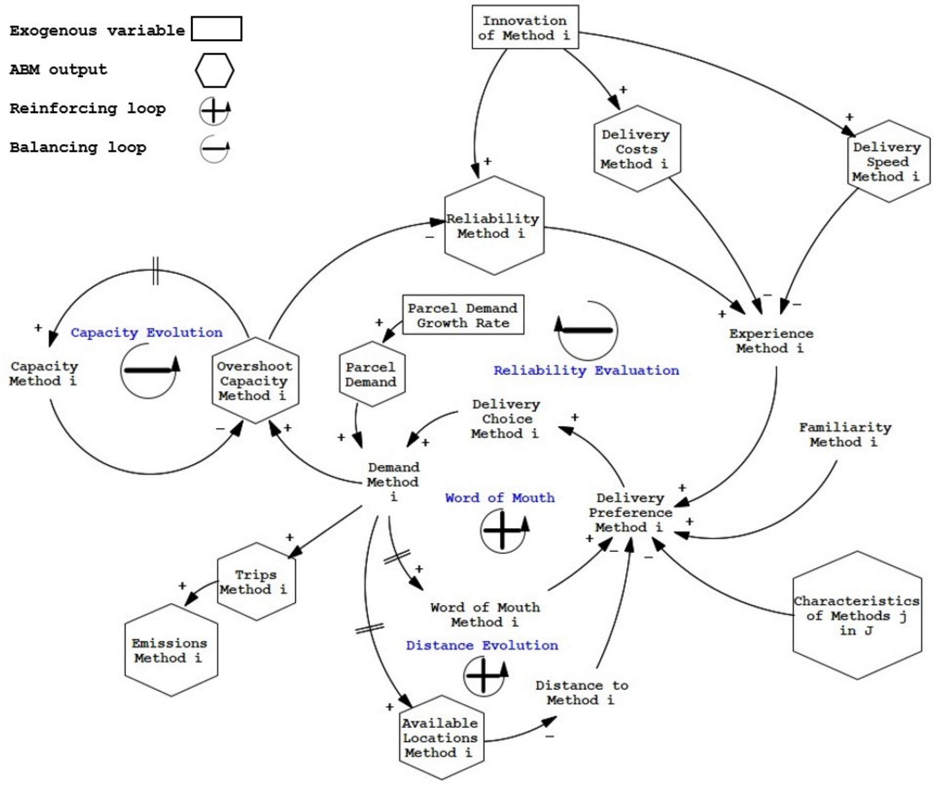


Fig. 2. The SD Causal Loop Diagram

model is a hybrid, the SD model receives input from the ABM. Each hexagon indicates an ABM output.

The centre of the causal loop diagram (CLD) is formed by *Delivery Preference Method i*. Which, according to the DOI theory, leads via *Delivery Choice Method i* to the demand of each delivery method. Preference for a delivery method is estimated on the operational performance, the WoM and familiarity effect. The operational performance is expressed by the reliability, costs, speed and pick-up distance specified per delivery method. The dynamic interactions of multiple loops (reinforcing (+) and balancing (-) with delays) lead to a complex system. Hence, a simulation is required to understand these interactions fully.

One of the main model steps within the SD model is the estimation of consumer preferences. Those preferences will be expressed as an aggregate probability distribution for the delivery methods. From that probability distribution, a preference will be randomly drawn for each parcel demand from the ABM.

To provide consumers feedback on their chosen delivery option, the characteristics of the performed delivery are estimated by calculating a performance score for each delivery method. This score can be seen as a quantitative representation of the perceived service quality of a consumer. The score of each delivery method i is calculated by:

$$S_i = \beta_s * \mu S_i + \beta_c * \mu C_i + \beta_r * \mu R_i + \beta_d * \mu D_i \quad \forall \quad i \in N \tag{1}$$

where μS_i is the average normalised delivery speed score, μC_i the average normalised delivery cost score, μR_i the average normalised reliability score, μD_i the average normalised distance score for the consumer to pick-up their parcel. The weights of the attributes are denoted by $\beta_s, \beta_c, \beta_r, \beta_d$. The value of the attribute levels and beta weights are based on previous studies [7] [8] [15] [30] [31] [32] [33] [34]; see VI and Table VII.

The delivery method preference probability is calculated with a logit model formulation, in which the performance score is subtracted from one, as the lowest score provides the best service; see Equation 2.

$$P_i(t+1) = \frac{e^{1-S_{it}}}{e^{\sum_{i=0}^{N} 1-S_{it}}} \quad \forall \quad i \in N \tag{2}$$

It is expected that self-collection and drone delivery methods will grow or decline due to the evolving preferences of consumers. Based on the previous demand-supply ratio, the distribution of self-collection points and the number of drones per depot can evolve. Van delivery and crowdshipping will not evolve because van delivery has an unlimited supply, and crowdshipping depends on the availability of crowdshippers, which is assumed to be constant.

To model WoM, the Bass diffusion model is used with one extension [23]: the performance of a delivery method determines the effect of WoM [35] [36]. Equation 3 shows the formula used in the model.

$$P_i(t+1) = P_i(t) + \frac{q\left(\frac{nP_i(t)}{\sum_{j=0}^{N} nP_i(t)}\right)\left(\left[\sum_{j=0}^{N} nP_i(t)\right] - nP_i(t)\right)\left(\frac{\sum_{i=0}^{N} S_{it}}{N} - S_{it}\right)}{\sum_{j=0}^{N} nP_i(t)} \quad \text{if } P_i(t) \neq 0 \quad \forall \quad i \in N \quad (3)$$

Here, q is the coefficient of imitation, and nP_i is the number of parcels of method i. Although not shown in this equation, the preference for each delivery method is divided by the sum of all preferences to maintain a summed total preference of 100%. Lastly, it could be possible that a delivery method is new and that, based on the potential performance, consumers will adopt it. In that case, the preference probability will be set to 3%, the percentage of innovators [23] [36].

The familiarity effect is expressed in Equation 4. It calculates the average chance that a consumer has never used a particular delivery method before and the resistance consumers thus have to use that delivery method [22] [24].

$$P_i(t+1) = \frac{P_i(t) * \left(1 - \left(\prod_{P_i t=0}^{t} 1 - P_i(t)\right) * \Omega_{chance}\right)}{\sum_{i=0}^{N}\left[P_i(t) * \left(1 - \left(\prod_{P_i t=0}^{t} 1 - P_i(t)\right) * \Omega_{chance}\right)\right]} \quad \forall \quad i \in N \quad (4)$$

D. The Agent-Based Model

The ABM is an existing agent-based model for urban freight transport: The Multi-Agent Simulation System for Goods Transport (MASS-GT). MASS-GT consists of three modules [37]: 1) a shipment synthesizer that replicates the strategic processes behind freight transportation demand; 2) a tour formation model on the tactical level which assigns shipments to tours and vehicles; 3) a network model where route choices are simulated. The parcel demand is estimated with an ordered logit model on an MPN survey with the explanatory variables: age, income and degree of urbanisation.

Parcel delivery with vans is already modelled in MASS-GT. Tour formation is based on capacity and time constraints, and tours originate and terminate at the same carrier depot [38]. A successful delivery rate of 75% is used [7]. The number of vans per depot is unconstrained.

In addition, MASS-GT has a module for crowdshipping for this study area in the version used for the HARMONY study [39]. For each crowdshipping request, the model allocates a crowdshipper that performs the delivery by van or bike [17], based on the crowdshippers utility if remuneration is provided for their work/effort. The consumer delivery fee is the remuneration plus a profit margin of 15%. The process ends when all parcels are allocated to the crowd or when crowdshippers with a higher utility for acting as a crowdshipper than for performing the regular trip are no longer available. Also, for crowdshipping, a delivery success rate of 75% is assumed.

The initial distribution of the self-collection points is based on the Dutch average number of inhabitants per self-collection point, and for each zone, the distance to the nearest self-collection point is calculated. Parcels for zones without a self-collection point are placed in the closest neighbour zone. If this demand exceeds the capacity, parcels will be randomly drawn until the capacity is reached. The remaining parcels are not delivered. Lastly, the assigned parcels for self-collection are assigned within the van tour formation.

Drones repeatedly fly from a depot to a consumer and back to the same depot. The flight distance is calculated by the Euclidean distance between the depot location and the centroid of the requested zone. A flight range of 10 kilometres and a flight speed of 45 km/h are assumed [40]. In addition, each flight will consist of a drone making two vertical movements to reach a safe flight height, a drop-off and reload time. The fee for drone delivery is €0.10 per kilometre [40], multiplied with a profit margin of 15%.

The demand for parcel deliveries by specific methods is converted into an assignment of delivery methods for each parcel. Not every parcel can or will be delivered with the preferred delivery method of the consumer because there could be no capacity, no crowdshipper can be found, or the delivery location is outside the drone range. In Fig. 3, a flow diagram is shown that represents the assignment procedure in this module. When the initial delivery cannot be performed successfully, a penalty is given to the reliability and speed score.

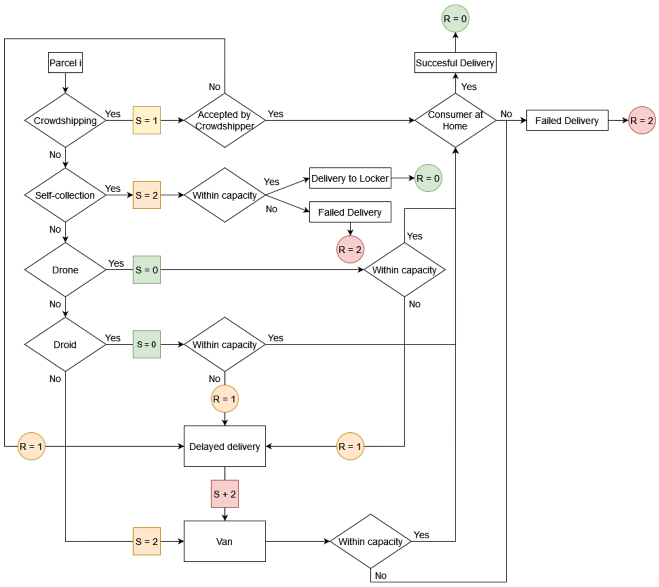


Fig. 3. Delivery assignment flow diagram with Speed Score (S) and Reliability Score (R).

Subsequently, the parcel carrier for each parcel is determined based on observed market shared [41]. After the assignment to a parcel carrier, the parcel is placed at the nearest depot of that carrier with respect to the delivery location. Then, trips will be scheduled that combine the delivery of parcels with the same mode. MASS-GT groups parcels on geographical proximity and provides a plausible set of tours [41]. After the cluster forming delivery tours are scheduled,

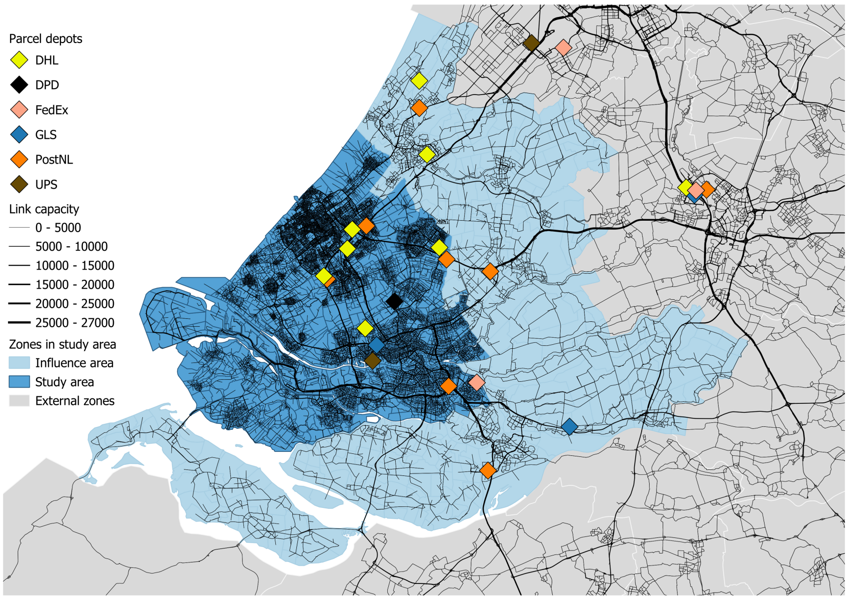


Fig. 4. The study area of MASS-GT with the road network, zonal distinction and courier depot locations.

hereby, the nearest neighbour approach is used to minimise the tour distance.

For van delivery, the scheduled tours are converted into trip matrices, which are assigned to a congested road network [37]. This assignment is based on an all-or-nothing shortest path assignment with generalised transportation costs of congested travel times. With those routes, the emissions for each used link are calculated using emission factors (g/km) [41]. A distinction in road type (urban, rural, highway) is made for these calculations and full and empty vehicles [42] with linear interpolation across the tour. The route calculations in MASS-GT estimate the route distance between the origin and destination zone. However, within each zone, multiple parcels can be delivered, for which additional kilometres must be driven. The in-zone drop-off location for each parcel request is unknown. Therefore, it is assumed that an additional distance is made by vans, which is twice the average Manhattan distance from the centroid to the average Euclidean intrazonal distance.

IV. Case Study

The developed simulation model will be applied in a case study on the province of South Holland. South Holland is a highly urbanised region containing multiple large cities and the seaport of Rotterdam; Fig. 4 shows a map of the area. The province has a population of 3.6 million, 50 municipalities and a surface area of 3.403 km^2. Both Rotterdam and the Hague are highly urbanized areas, yet, most other zones are still quite densely populated and cannot be called rural. South Holland has a well-developed transport network with various high-capacity roads and highways. In total, there are 29 carrier depots in and near South Holland.

A. Simulation Scenarios

With the presented model, the parcel freight logistics of South Holland will be simulated for three different scenarios with a simulation horizon of five years. Each iteration inside a simulation run represents a time step of a quarter of a year. Thus each simulation will consist of 20 iterations to represent the evolution over time. A constant yearly growth rate of 21.6% for the parcel demand will be implemented [43], while the population and its distribution will be considered constant. The following delivery scenarios will be simulated, with the corresponding initial distribution of the delivery method preferences in Table I [11] [30]:

1) Current State: van and self-collection delivery. The self-collection points can evolve in number and capacity.
2) Crowdshipping: As crowdshipping has to overcome fewer regulatory and technological barriers than drones, it is expected that crowdshipping will be the first delivery innovation that will be implemented in the Netherlands. This scenario will simulate the coexistence of van, self-collection and crowdshipping in the study area.
3) Full innovation: Drones are added to the crowdshipping scenario.

TABLE I
INITIAL DISTRIBUTION OF DELIVERY METHOD PREFERENCE

Delivery Method	**Initial Preference Scenario 1**	**2**	**3**
Van	80%	59%	59%
Self-collection	20%	20%	20%
Crowdshipping		21%	21%
Drone			0%

B. Validation and Verfication

To verify and validate the proposed approach, three tests are performed: 1) Test for face validity; 2) Test for various input parameters; 3) Compare model predictions with the performance of the actual system or with predictions from other studies [44].

A thorough sensitivity analysis is performed for various estimated model parameters. In all cases, irrespective of the sensitivity, the model output changes in a logical manner that corresponds with the theory. Only in the case of an extreme value for the beta of reliability the model showed unstable behaviour; see the examples of Fig. 9 and 10. The preference probability of all delivery methods becomes very unstable. This is caused by the highly alternating performance scores of self-collection and drone delivery. Both performance scores fluctuate strongly as the demand overshoots the capacity in one run. Subsequently, the preference shifts to other delivery methods, which results in a high performance of self-collection and drone. Hence, the demand rises and overshoots the capacity. As all results are explainable and the developed model functions appropriately, the model can be used to evaluate the interactions within the system. Furthermore, a time horizon test is carried out, which also confirms that the model behaviour is consistent and explainable.

To validate the model, the model output is compared with real-world results and other research. In Table II, the market shares, as estimated by the model and as presented in literature are compared. The literature values for van delivery are the currently operated percentage in the Netherlands. The market shares for self-collection, crowdshipping and drone are estimation results based on stated preference studies. Because the model results are, to a large extent, in line with other research, it can be concluded that the model behaviour does reflect the real-world system and can thus be used to analyse the last-mile system.

TABLE II
MARKET SHARE MODEL ESTIMATION VERSUS LITERATURE

Market Share	**Scenario t=20** [quarter] 1	2	3	**Literature Value**	**Scource**
Van	69%	53%	39%	79.5 - 81.5%	[43] [45]
Self-collection	31%	25%	18%	18 - 29%	[30] [46]
Crowdshipping		22%	15%	21 - 27%	[11]
Drone			28%	7.18 - 53%	[47] [48]

C. Results

Table III presents the results of the three scenario simulations. For each scenario, the simulation is performed five times. The averages of those five runs are presented. In all three scenarios, the stochasticity in the model produces very limited differences between the runs. The largest variance in preference probability occurs at $t = 9$ with drone delivery; $\sigma = 2.4E{-}5$.

The results in Table III indicate that most consumers prefer van delivery in all three scenarios. The performance score of van delivery is constantly low, which represents a high operational service quality; for example, see Fig. 6. In the iterations near the end of the time horizon, the WoM effect is the main driving force for consumers to choose this option; as an example, see Fig. 11. Due to the growth in parcel demand, this leads to an enormous increase in the number of trips that have to be performed per day by van.

The introduction of other delivery methods can significantly reduce the market share of van delivery. Even though the performance of the other delivery methods is, on average, of lower quality, many consumers will choose those delivery methods. In Fig. 5, the evolving consumer preference distribution can be seen for scenario 3, and Fig. 6 shows that the performance score of van delivery is, on average, the best after a number of iterations.

A clear illustration of the interaction between demand and supply occurs with self-collection. Over time the supply of self-collection points reduces, see Table III, which can be explained by the demand being below the removal threshold in many zones. This results in a worsening performance score, see Fig. 6, where the reliability mainly decreases due to the number of parcels not fitting in the shrunken capacity. This creates a negative feedback loop between consumer demand and the supply of self-collection. Still, around one-fifth of the consumers are expected to choose self-collection. The number of trips drops during the simulation because parcels can be grouped and delivered at fewer self-collection points. The spatial results show that self-collection can be a competitive delivery method in dense urban areas; see Fig. 8.

TABLE III
KPIS SCENARIOS

Indicator	**Method**	**Scenario t=0** [quarter] 1	2	3	**Scenario t=20** [quarter] 1	2	3
Market share	Van	80%	59%	59%	69%	53%	39%
	Self-Collect	20%	20%	20%	31%	25%	18%
	Crowdshipping		21%	21%		22%	15%
	Drone			0%			28%
Trips per day	Van	92485	66815	66716	153277	116497	84181
	Self-Collect	11343	11171	11171	8052	6890	5053
	Crowdshipping		12079	12237		23169	17729
	Drone			0			36324
Vehicle kilometers per day	Van	90394	82866	82273	108793	96017	83205
	Crowdshipping		26492	26702		45590	36785
	Drone			0			195858
CO2 emissions per day	Van	18484	16646	16531	22904	19916	16940
	Crowdshipping		1782	1792		3047	2462
	Drone			0			0
Number of Self-Collection points		3012	3012	3012	2250	1956	1508
Total capacity Self-Collection points		60240	60240	60240	44500	39000	30000
Number of Drones				29			994

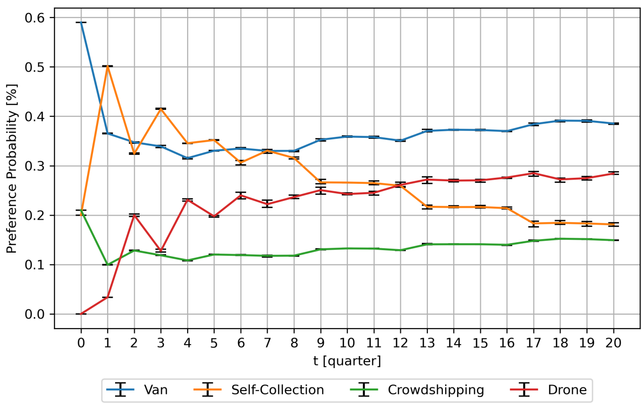


Fig. 5. Preference probability in scenario 3.

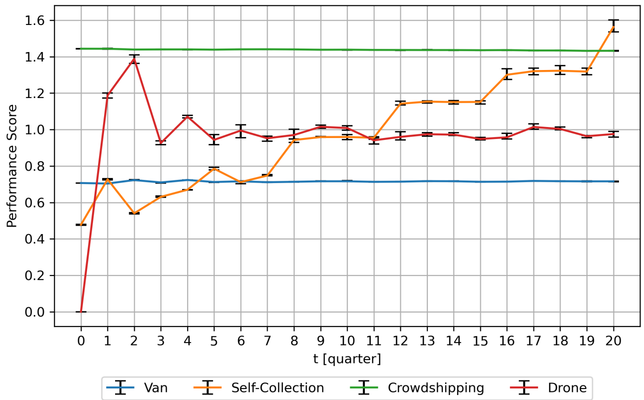


Fig. 6. Performance score in scenario 3.

The results show that with the introduction of more delivery methods, fewer consumers will prefer self-collection; see Table III. Because of those lower demands, the distribution of self-collection becomes scarcer, and the total capacity at $t = 20$, after 5 years, is around 13% and 31% lower for scenarios 2 and 3 than in scenario 1, respectively. Consequently, the performance score of self-collection in scenario 3 evolves to a slightly worse level than in scenarios 1 and 2.

Crowdshipping has, like van delivery, a constant performance score; see Fig. 6. Additionally, the performance is very consistent across all zones. Because of the stable performance, the evolution of crowdshipping is very comparable between scenarios 2 and 3. However, in scenario 3, the competition with drone delivery results in a lower market share.

The introduction of drone delivery has a large impact on the predicted market shares of the other delivery methods, as can be seen in Table III by the reduced market shares of the other methods compared with scenarios 1 and 2. Spatial results show that drones do perform well around the depots. However, the performance in the outskirts of the study area is generally low. The development of the flight range makes drone delivery accessible to more consumers. Still, the average performance score does not improve; see Fig. 6. This is because the high accessibility effect is cancelled out by the increased delivery costs, which are linked to the flight distance.

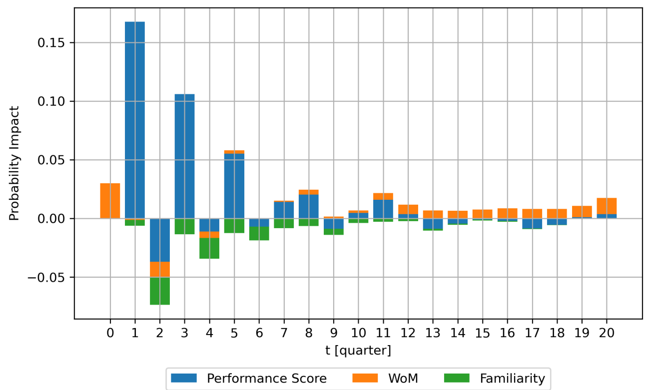


Fig. 7. Impact on the preference probability for drone in scenario 3.

To fulfil the demand, drones must travel large amounts of vehicle kilometres. In just five years, the drone fleet has to grow from 29 drones in total to almost 1000 in order to fulfil the demand. Consequently, there will be four depots that must operate more than 100 drones, with a maximum of 183.

Consumers evolve their preferences because of the performance score, WoM and familiarity. In Fig. 7, the impact of these three factors on the preference probability for drone is shown for consecutive iterations. Initially, the only influence comes from the WoM effect because that factor accounts for the adoption of innovators. Based on the operational performance that those innovators experience, the logit model assumes a strong uptake in users after the first quarter, $t = 1$. At that iteration, the magnitude of WoM is very limited due to the small group of consumers that can spread WoM. On the contrary, the familiarity effect is relatively strong at the start because only a few consumers have used drones, and thus many consumers still need to familiarize themselves with drones. When the performance score of drone delivery stabilizes, from Fig. 6 around $t = 10$, also the impacts become smaller. In the end, the WoM effect becomes the most important reason for the continued growth of the market share of drones.

Finally, the large yearly growth in parcel demand results in higher vehicle kilometres and CO2 emissions, regardless of the scenario. As can be seen in Table IV, the total demand almost doubled in five years. The introduction of crowdshipping reduces the demand for van delivery and self-collection, yet it will result in more vehicle kilometres that need to be made on the same road infrastructure as van delivery. However, crowdshipping produces less CO2 per km, and therefore the CO2 emissions are very comparable with scenario 1. The high CO2 emissions in scenario 2 can be explained by vans becoming more efficient when the demand is higher. With increasing parcel demand, the number of zones that need to be visited increases slightly, and those transports contribute to the largest share of vehicle kilometres and emissions. Hence, the positive effect of the substitution of van and self-collection delivery by crowdshipping is partly reversed. This can also be seen in Table V, where the average kg CO2 emission per parcel is shown. In all scenarios, this number improves. The main

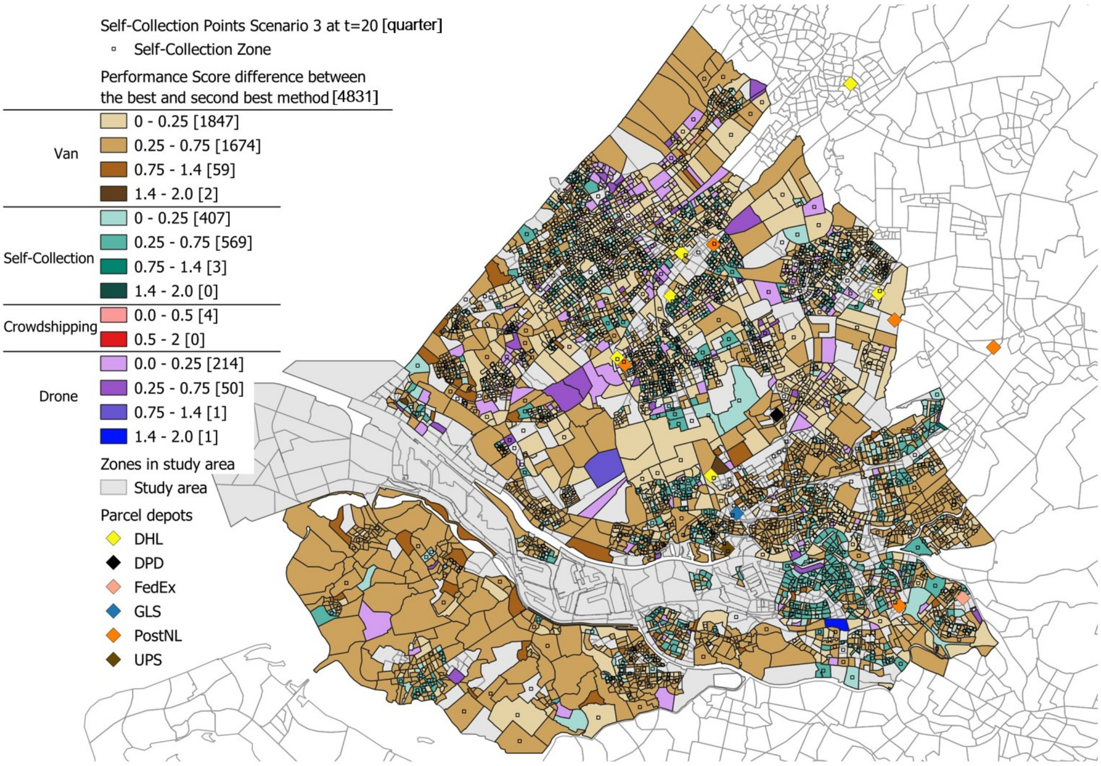


Fig. 8. Zonal distribution of the relative performance of all delivery methods in scenario 3 after 20 quarters. A zone is filled with the colour of the best-performing delivery method. The mark is the difference between the performance score of the best and second-best delivery method.

reason herefore is vans becoming more efficient due to the absolute increase in parcel demand delivery by van. Scenario 3 results in the lowest CO2 emissions due to drone delivery being emission-free and results in 15% less CO2 emission than scenarios 1 and 2. Additionally, the van vehicle kilometres and emissions within zones are much lower than in the other scenarios, which reduces the last-mile burden in urbanised areas and on low-capacity roads.

TABLE IV
SCENARIO COMPARISON X1000

Indicator	**Scenario t=0** [quarter] **1**	**2**	**3**	**Scenario t=20** [quarter] **1**	**2**	**3**
Parcel Demand [parcels]	116	116	116	223	223	223
Vehicle Kilometres	90.4	109	109	109	142	316
CO2 Emissions [kg]	18.5	18.4	18.3	22.9	23.0	19.4

TABLE V
AVERAGE C02 EMISSIONS PER PARCEL [KG/PARCEL]

Time [quarter]	**Scenario 1**	**Scenario 2**	**Scenario 3**
t = 0	0.160	0.159	0.159
t = 20	0.103	0.103	0.087

V. CONCLUSION

In this research, a hybrid simulation model is developed that estimates consumer preferences and the operation of various delivery methods for consecutive interactions. An SD model simulates the preference evolution of consumers over time, and an ABM simulates the carriers delivering the parcel from the depot to the receiver and the evolution of self-collection and drone capacity. The presented model is able to empirically simulate consumer behaviour influenced by the synthetic delivery experience, WoM and the familiarity effect.

The contribution of this new model is threefold. Firstly, by simulating the interaction between the demand and supply of several parcel delivery methods in a spatial model, it is possible to model policy measures and delivery operation strategies on an aggregate to zonal level. This approach allows the exploration of the complex last-mile system where multiple delivery methods compete and complement each other.

The second contribution is the recognition of stochasticity in the delivery operation. The simulations estimate a delivery method's reliability at various demand and supply levels. This supports future research into the interaction between consumers and the reliability of delivery methods.

Thirdly, this modelling approach provides insight into the direction and magnitude of various factors that take place, with the added novelty of gathering that empirical data at multiple time points. This is a valuable model feature, as stated preference and especially revealed preference research cannot determine such data.

From the performed case study on the province of South Holland, several recommendations for policymakers can be established. If policymakers or parcel carriers want to reduce the demand for van delivery, providing additional delivery methods could lessen that demand. However, those delivery methods should establish very high service levels, which is not likely or costly in many zones. A significant impact could also be organised by reducing the service of van delivery, mainly by introducing a delivery fee, as consumers are cost-sensitive. Furthermore, it is concluded that self-collection is a highly competitive delivery method in dense urban areas. To make it even more successful, municipalities should direct or build a vast network, especially of white-label, automated lockers. Finally, as both crowdshipping and drone delivery are new innovations, policymakers have the opportunity to steer regulations such that societal benefits are prioritised.

VI. Discussion and Future Work

The simulation model and results can be improved by adding or applying various advancements. First of all, by implementing the SD model part at a lower level, for example, a neighbourhood or municipality, instead of the aggregate level of the entire study area, a more precise estimation can be given of where particular delivery methods prevail or are irrelevant. In addition, this development allows zonal differentiation in, for example, attribute weights. If a sufficient data source is available, the dynamics within the SD model can be implemented in an ABM, which offers the opportunity to model at the individual consumer level, allowing for even more heterogeneity. Secondly, a moving memory can be considered, especially for longer simulation horizons. Thirdly, a start-up period for the model can be introduced, as the consumer preferences fluctuate to a stable situation in the first two to four iterations. Furthermore, parcel characteristics like weight and size are disregarded, which overestimates the possibility of especially drone delivery.

Besides these model improvements, it is recommended that future research and data collection efforts are carried out on consumer preferences for last-mile delivery. Limited data is currently available, and general best practices have not been settled. Additionally, future research could implement non-performance-related delivery attributes to describe consumer preferences.

References

[1] S. Punte, L. Tavasszy, A. Baeyens, and F. Liesa, "A framework and process for the development of a roadmap towards zero emissions logistics 2050," Alliance for Logistics Innovation through Collaboration in Europe, 12 2019.

[2] ITF, "The freight space race: Curbing the impact of freight deliveries in cities," *International Transport Forum Policy Papers*, vol. 109, 2022.

[3] N. Boysen, S. Fedtke, and S. Schwerdfeger, "Last-mile delivery concepts: a survey from an operational research perspective," *Or Spectrum*, vol. 43, pp. 1–58, 2021.

[4] "Self-driving vehicles in logistics. a dhl perspective on implications and use cases for the logistics industry," 2014, dHL Trend Research.

[5] M. Joerss, J. Shcröder, F. Neuhaus, C. Klink, and F. Mann, "Parcel delivery: the future of last mile," McKinsey&Company, 9 2016.

[6] S. B. Hepp, "Innovation in last mile delivery: meeting evolving customer demands: the case of in-car delivery," Ph.D. dissertation, 2018.

[7] H. Buldeo Rai, S. Verlinde, and C. Macharis, "The "next day, free delivery" myth unravelled: Possibilities for sustainable last mile transport in an omnichannel environment," *International Journal of Retail & Distribution Management*, vol. 47, no. 1, pp. 39–54, 2019.

[8] I. Maltese, M. Le Pira, E. Marcucci, V. Gatta, and C. Evangelinos, "Grocery or@ grocery: A stated preference investigation in rome and milan," *Research in Transportation Economics*, vol. 87, p. 101096, 2021.

[9] B. Ignat and S. Chankov, "Do e-commerce customers change their preferred last-mile delivery based on its sustainability impact?" *The International Journal of Logistics Management*, vol. 31, no. 3, pp. 521–548, 2020.

[10] B. Asdecker, "How dare you replace my deliveryman?! acceptance of last-mile transportation innovations–a qualitative perspective," 2020.

[11] H. Buldeo Rai, S. Verlinde, and C. Macharis, "Who is interested in a crowdsourced last mile? a segmentation of attitudinal profiles," *Travel Behaviour and Society*, vol. 22, pp. 22–31, 2021.

[12] ACM, "Marktstudie last mile pakketbezorging," De Autoriteit Consument en Markt, 5 2020, rapport ACM/UIT/534379.

[13] S. Leon, C. Chen, and A. Ratcliffe, "Consumers' perceptions of last mile drone delivery," *International Journal of Logistics Research and Applications*, vol. 26, no. 3, pp. 345–364, 2023.

[14] W. A. Mohammad, Y. Nazih Diab, A. Elomri, and C. Triki, "Innovative solutions in last mile delivery: concepts, practices, challenges, and future directions," in *Supply Chain Forum: An International Journal*. Taylor & Francis, 2023, pp. 1–19.

[15] V. Gatta, E. Marcucci, M. Nigro, S. M. Patella, and S. Serafini, "Public transport-based crowdshipping for sustainable city logistics: Assessing economic and environmental impacts," *Sustainability*, vol. 11, no. 1, p. 145, 2018.

[16] A. Punel, A. Ermagun, and A. Stathopoulos, "Studying determinants of crowd-shipping use," *Travel Behaviour and Society*, vol. 12, pp. 30–40, 2018.

[17] R. J. Tapia, I. Kourounioti, S. Thoen, M. de Bok, and L. Tavasszy, "A disaggregate model of passenger-freight matching in crowdshipping services," *Transportation Research Part A: Policy and Practice*, vol. 169, p. 103587, 2023.

[18] A. Shabbir, S. A. Malik, and S. Y. Janjua, "Equating the expected and perceived service quality: A comparison between public and private healthcare service providers," *International Journal of Quality & Reliability Management*, 2017.

[19] X. Wang, K. F. Yuen, Y. D. Wong, and C.-C. Teo, "E-consumer adoption of innovative last-mile logistics services: A comparison of behavioural models," *Total Quality Management & Business Excellence*, vol. 31, no. 11-12, pp. 1381–1407, 2020.

[20] S. J. Grawe, "Logistics innovation: a literature-based conceptual framework," *The International Journal of Logistics Management*, vol. 20, no. 3, pp. 360–377, 2009.

[21] G. De La Torre, T. Gruchmann, V. Kamath, A. Melkonyan, and K. Krumme, "A system dynamics-based simulation model to analyze consumers' behavior based on participatory systems mapping–a "last mile" perspective," *Innovative Logistics Services and Sustainable Lifestyles: Interdependencies, Transformation Strategies and Decision Making*, pp. 165–194, 2019.

[22] Y. Vakulenko, P. Shams, D. Hellström, and K. Hjort, "Service innovation in e-commerce last mile delivery: Mapping the e-customer journey," *Journal of Business Research*, vol. 101, pp. 461–468, 2019.

[23] F. M. Bass, "A new product growth for model consumer durables," *Management science*, vol. 50, no. 12 supplement, pp. 1825–1832, 2004.

[24] C. Fandos Herrera and C. Flavián Blanco, "Consequences of consumer trust in pdo food products: the role of familiarity," *Journal of Product & Brand Management*, vol. 20, no. 4, pp. 282–296, 2011.

[25] S. Shepherd, "A review of system dynamics models applied in transportation," *Transportmetrica B: Transport Dynamics*, vol. 2, no. 2, pp. 83–105, 2014.

[26] E. Pruyt, "Small system dynamics models for big issues: Triple jump towards real-world complexity," 2013.

[27] C. Thaller, U. Clausen, and R. Kampmann, "System dynamics based, microscopic freight transport simulation for urban areas: A methodological approach," in *Commercial Transport: Proceedings of the 2nd Interdiciplinary Conference on Production Logistics and Traffic 2015*. Springer, 2016, pp. 55–72.

[28] B. J. Angerhofer and M. C. Angelides, "System dynamics modelling in supply chain management: research review," in *2000 Winter Simulation Conference Proceedings (Cat. No. 00CH37165)*, vol. 1. IEEE, 2000, pp. 342–351.

[29] R. Martin and M. Schlüter, "Combining system dynamics and agent-based modeling to analyze social-ecological interactions—an example from modeling restoration of a shallow lake," *Frontiers in Environmental Science*, vol. 3, p. 66, 2015.

[30] E. Molin, M. Kosicki, and R. van Duin, "Consumer preferences for parcel delivery methods: The potential of parcel locker use in the netherlands," *Eur. J. Transp. Infrastruct. Res*, vol. 22, pp. 183–200, 2022.

[31] E. Caspersen and S. Navrud, "The sharing economy and consumer preferences for environmentally sustainable last mile deliveries," *Transportation Research Part D: Transport and Environment*, vol. 95, p. 102863, 2021.
[32] D. H. Nguyen, S. De Leeuw, W. Dullaert, and B. P. Foubert, "What is the right delivery option for you? consumer preferences for delivery attributes in online retailing," *Journal of Business Logistics*, vol. 40, no. 4, pp. 299–321, 2019.
[33] L. K. de Oliveira, E. Morganti, L. Dablanc, and R. L. M. de Oliveira, "Analysis of the potential demand of automated delivery stations for e-commerce deliveries in belo horizonte, brazil," *Research in Transportation Economics*, vol. 65, pp. 34–43, 2017.
[34] V. Gatta, E. Marcucci, M. Nigro, and S. Serafini, "Sustainable urban freight transport adopting public transport-based crowdshipping for b2c deliveries," *European Transport Research Review*, vol. 11, no. 1, pp. 1–14, 2019.
[35] R. Wong and S. Y. Sheng, "A business application of the system dynamics approach: word-of-mouth and its effect in an online environment," *Technology Innovation Management Review*, vol. 2, no. 6, 2012.
[36] V. Mahajan, E. Muller, and F. M. Bass, "New product diffusion models in marketing: A review and directions for research," *Journal of marketing*, vol. 54, no. 1, pp. 1–26, 1990.
[37] M. de Bok, L. Tavasszy, and S. Thoen, "Application of an empirical multi-agent model for urban goods transport to analyze impacts of zero emission zones in the netherlands," *Transport Policy*, vol. 124, pp. 119–127, 2022.
[38] S. Thoen, L. Tavasszy, M. de Bok, G. Correia, and R. van Duin, "Descriptive modeling of freight tour formation: A shipment-based approach," *Transportation Research Part E: Logistics and Transportation Review*, vol. 140, p. 101989, 2020.
[39] M. de Bok, L. Tavasszy, I. Kourounioti, S. Thoen, L. Eggers, V. Nielsen, and J. Streng, "Application of the harmony tactical freight simulator to a case study for zero emission zones in rotterdam," in *Transportation Research Board Annual Meeting*. Transportation Research Board (TRB), 2021, transportation Research Board Annual Meeting ; Conference date: 05-01-2021 Through 29-01-2021. [Online]. Available: http://www.trb.org/AnnualMeeting/AnnualMeeting.aspx
[40] R. D'Andrea, "Guest editorial can drones deliver?" *IEEE Transactions on Automation Science and Engineering*, vol. 11, no. 3, pp. 647–648, 2014.
[41] S. Thoen, M. de Bok, L. Eggers, and I. Kourounioti, "E-commerce in strategic freight models: an implementation in the province of zuid-holland," 2021, slides for European Transport Conference 2021.
[42] S. Thoen, M. de Bok, and L. Tavasszy, "Shipment-based urban freight emission calculation," in *2020 Forum on Integrated and Sustainable Transportation Systems (FISTS)*. IEEE, 2020, pp. 372–377.
[43] ACM, "Marktstudie last mile pakketbezorging," De Autoriteit Consument en Markt, 6 2022.
[44] R. G. Sargent, "Verification and validation of simulation models," in *Proceedings of the 2010 winter simulation conference*. IEEE, 2010, pp. 166–183.
[45] ACM, "Post- en pakkettenmonitor 2021," De Autoriteit Consument en Markt, 2021, rapport ACM/UIT/579837.
[46] P. Buijs and R. Niemeijer, "A greener last mile: Reviewing the carbon emission impact of pickup points in last-mile parcel delivery," *Available at SSRN 4169737*, 2022.
[47] S. H. Kim, "Choice model based analysis of consumer preference for drone delivery service," *Journal of Air Transport Management*, vol. 84, p. 101785, 2020.
[48] R. Merkert, M. C. Bliemer, and M. Fayyaz, "Consumer preferences for innovative and traditional last-mile parcel delivery," *International Journal of Physical Distribution & Logistics Management*, 2022.

Appendix

TABLE VI
ATTRIBUTE BETA WEIGHTS

Attribute	Weight
β_{Speed}	1
β_{Costs}	2
$\beta_{Reliability}$	2
$\beta_{pick-upDistance}$	1

TABLE VII
ATTRIBUTE LEVELS

Attribute	Levels	Description	Normalised Performance Score
Delivery Speed	0	several hours	0
	1	same-day	0.04
	2	1-day	0.2
	3	2-days or more	1
Delivery Costs	0	free delivery	0
	1	costs $<$ 2 euro	0.33
	2	costs $<$ 5 euro	0.66
	3	costs $>$ 5 euro	1
Delivery Reliability	0	as expected	0
	1	delayed delivery and/or other delivery method	0.33
	2	unsuccessful delivery	1
Pick-up Distance	0	distance $<$ 100 meters	0
	1	distance $<$ 300 meters	0.387
	2	distance $<$ 500 meters	0.613
	3	distance $<$ 1000 meters	0.774
	4	distance $<$ 2000 meters	0.898
	5	distance $>$ 2000 meters	1

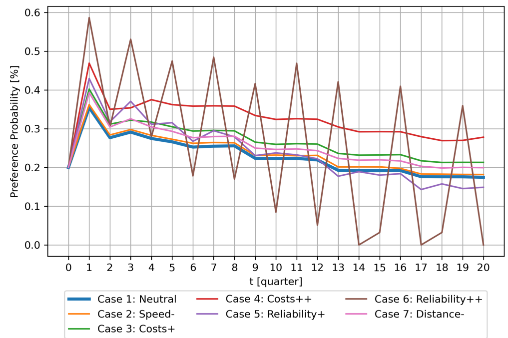


Fig. 9. Probability self-collection with different beta weights.

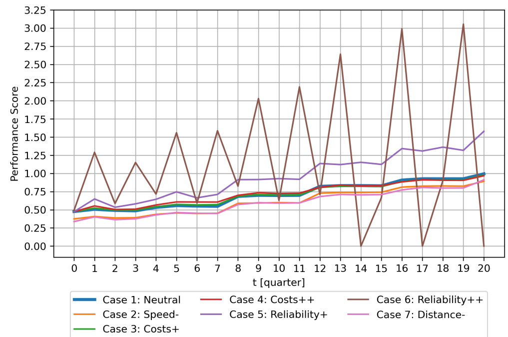


Fig. 10. Peformance score self-collection with different beta weights.

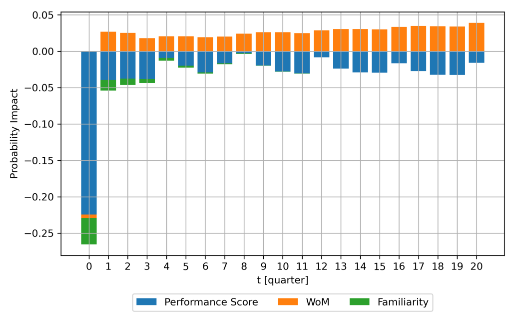


Fig. 11. Impact on the preference probability of van in scenario 3.