Model-Based Decision Support on the Last Mile of Distribution-Logistics

Dissertation zur Erlangung des Grades eines Doktors der Wirtschaftswissenschaft

eingereicht an der Fakultät für Wirtschaftswissenschaften der Universität Regensburg

vorgelegt von: Christian Brabänder

Berichterstatter:
Prof. Dr. Andreas Otto
Prof. Dr. Guido Schryen

Tag der Disputation: 06.05.2019

"Science's models are not true, and that's exactly what makes them useful. They tell simple stories that our minds can grasp. They are lies-to-children, simplified teaching stories, and none the worse for that. The progress of science consists of telling ever more convincing lies to ever more sophisticated children."

T. Pratchett, I. Stewart, J. Cohen (1999), The Science of Discworld

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AAMBR	Axis-aligned minimum bounding rectangle
API	Application Programming Interface
AVNS	Adaptive Variable Neighboorhood Search
BBP	Break-bulk point
BI	Best improvement/steepest decent
BKOV	Best known objective value
CEP	Courier, Express, and Parcel
CG	Consumer goods
CoDi	Collection and Delivery
CT	Computation time
CTP	Collaborative Transportation Planning
CWH	Central warehouse
DDP	Distribution Districting Problem
DRP	Distribution requirement planning
DSD	Direct Store Delivery
DSS	Decision Support System
FI	First improvement/first descent
FMCG	Fast Moving Consumer Goods
FTL	Full truckload
GF	Groupage freight
GHG	Greenhouse gas
GIS	Geographical Information System
HVRP	Heterogeneous VRP
Imp. CT	Relative computation time improvement
IP	Integer Program
IOTP	Integrated Operational Transportation Planning
ITPP	Integrated Transportation Planning Problem
JIT	Just-in-Time
LKH	Lin-Kernighan-Helsgaun
LSP	Logistics service provider
LTL	Less than truckload
MDVRPPC	Multiple depots and called their problem multi-depot vehicle routing problem with private fleet and common carriers
MIP	Mixed Integer Program
MoB	Make or Buy
MTZ	Miller-Tucker-Zemlin
OR	Operations Research
OV	Objective value
PCMDVRPNL	Prize-Collecting Multi-Depot Vehicle Routing Problem with Non-Linear costs
PCTSP	Prize Collecting Traveling Salesman Problem
PDP	Pick-up and delivery problem
POS	Point-of-Sale
RD	Route dividing
RIP	Randomized construction Improvement-Perturbation
RQ	Research question
RVND	Randomized Variable Neighborhood Descent
RWH	Regional warehouse
SA	Simulated Annealing
SCM	Supply Chain Management
SM	Solution method
SMCG	Slow Moving Consumer Goods
sVRP	Selective vehicle routing problem
TC	Total cost
TMS	Transport Management System
TS	Tabu search
TS+	Tabu search heuristic by ejection chains
TSP	Travelling Salesman Problem

VMI	Vendor Managed Inventory
VND	Variable Neighborhood Descent
VNS	Variable Neighborhood Search
VNSac	VNS using an acceptance criterion
VRFP	Vehicle Routing and Forwarding Problem
VRP	Vehicle Routing Problem
VRPPC	Vehicle routing problem with private fleet and common carriers
VRPPCdR	Vehicle routing problem with private fleet, multiple common carriers offering volume discounts, and rental options
XD	Cross dock

1 Introduction

Distribution is an important driver of success for shipping, freight-forwarding, and retailing companies, as well as for their consumers (Fuller et al. 1993; Lambert and Burduroglu 2000; Kumar 2001). The distribution of goods drives both the fulfilment cost and customer experience, hence also profitability (Chopra 2003, p. 123). Measuring the effects of distribution does not only restrain to cost and revenue accounting. All relevant impacts are measured, such as service level, speed and flexibility, ecological, and social aspects. The value of distribution originates from making a product available for further production, retailing, or its consumption. It is worthwhile but nevertheless difficult to convert the non-monetary effects into monetary units. Although it is impossible to value every capability, every soft driver, and every percent of service by the exact monetary value, the direction of the effect of almost all activities, decisions, and their ceteris paribus effects can be estimated (Lambert and Burduroglu 2000; Chopra 2003). The "last mile" of distribution has the reputation of being difficult, costly, and critical for success: the last mile is the hardest mile. The introductory section of this dissertation outlines distribution as a major research topic and focuses on that last mile. Sub-section 1.1 summarizes three perspectives on distribution. In 1.2, the essentials of distribution-logistics with a focus on consumer goods are discussed. Sub-section 1.3 deduces a definition of what the last mile is and emphasizes the issues and the decision problems of the last mile. In sub-section 1.4, the theoretical foundations of model-based decision support are laid and three selected focuses of model-based decision support on the last mile are outlined.

1.1 On perspectives of distribution

Distribution describes all activities of economic and physical transportation of goods (Raffée 1974, p. 195). The goals of distribution are "the seven rights": having the right goods, in the right amount, in the right quality, at the right place, at the right time, for the right customer, at the right price (Plowman 1964). In literature, there are at least three fields of research investigating distribution: Marketing, Logistics, and Supply Chain Management (SCM). In a nutshell, these perspectives view distribution like this:

- **Marketing** views distribution as the selection, design, and coordination of channels that bridge the gap between production and consumption.
- **Logistics** views distribution as the generic process that connects productive and consumptive processes by physical flows with the objective to create availability.
- **SCM** views distribution as a subsystem of the supply chain system[1], which performs all processes that relate a company's production with its customers' procurement processes.

[1] Herein, a system is understood as the whole of structure elements and their (potential) relations. I take the general definition of systems by Ulrich 1968, pp. 105–118 and agree with him that managerial decision-making requires a special understanding of systems as open, planned, and purposeful in order to enable system design. Ulrich elaborates on his definition that there must exist a structure of its elements given by their relations. Further, there may be a processual structure as events have a sequential structure over time. The function (= purpose) of a system is not part of the general, but of the special definition, of systems. What is meant is that there are natural systems without a purpose. Nevertheless, artificial systems like distribution do have functions.

Taking these perspectives is worthwhile, because changing the perspectives enables decision-makers to identify different decision-objects and leverage different decision-alternatives. If a decision-maker views distribution from yet another perspective, different questions come to mind and different decision-making processes are spotted (Kirsch 1998; Otto and Kotzab 2003). The "seven rights" are extremely difficult to achieve taking only one isolated perspective.

Example 1: The problem of supplying the right good to the right customer may be seen as a marketing issue: identify the customers' need and supply the right good accordingly. However, this relation benefits from a multi-perspective approach. In a situation of supply shortfall, logistics provides decision support models, how to allocate shortages, select back-orders, and ways how to replenish.

Example 2: The right time, place, amount, and quality are the function of distribution-logistics (cf. Ihde 2001). Under the implicit assumption of knowing what the customer wants, logistics supports decision-making on ways, how to get products there in time, on shipment sizes, and on assortments. Nevertheless, service-levels – availability of goods at the Point-of-Sale (POS) – of almost 100% are extremely expensive. Therefore, logistics has the goal of minimizing the total cost of shortages and overstocking: the marketing perspective is helpful to estimate the cost of shortages (Ihde 2001, p. 312).

Besides the already named, there are other perspectives imaginable[2]. However, it is believed that these three are the most important in terms of literature quantity and practical relevance. The following sub-sections 1.1.1-1.1.3 explore the outlined perspectives on distribution in a chronological order. Marketing has been there first and distribution was an all marketing topic. Then academics separated the acquisition of demand and the physical distribution of goods: logistics (Converse 1954; Drucker 1962; Alewell 1968). SCM is a rather young field of research coming from practitioners and consultants. Is has been named in the 1980's and academics started discussing SCM in the 1990's (Cooper et al. 1997).

1.1.1 Marketing: channel design

From a marketing perspective, distribution is viewed as an enabler which enables "the seven rights" – raising questions, like "What does our customer demand/need?" meaning "find the right supply" and "Who are our right customers?" meaning "find the right demand". Therefore, marketing raises questions at the upstream/supplier and the downstream/customer side of a distribution channel. Manufacturers and customers are connected by a channel (="pipeline") (Otto and Kotzab 2003; Coughlan et al. 2006). This channel bridges the gap between the production and consumption of goods. In a broader sense, Ahlert (2005) adds all activities, which are performed on real or immaterial goods, to distribution, in-between the institutional production and consumption of those

[2] From my point of view, the limitation to these three perspectives has two reasons. First, some perspectives are strongly linked to these three, such that a sharp differentiation is not helpful. For example, one could argue that organization theory provides helpful insights into distribution. Nevertheless, the relevant lessons from organization theory can be learned from the SCM perspective (Otto and Kotzab 2003). Second, many authors would probably see methodological perspectives such as operations research (OR) or statistics as the go-to field of research when thinking about distribution. Herein, methodology is seen as an (undeniably important) toolbox including tools that are helpful for many issues of distribution, but that is not a separate perspective. Especially the field of logistics applies OR techniques extensively.

goods. Thus, distribution comprises all activities that enable a purchase agreement. Ahlert includes all kinds of marketable offers in his definition, including goods, real estate, property rights[3], and services. From today's perspective, one should add digital products to that list. Converse (1954) acknowledges that there are the two halves of marketing: first, an immaterial half of creating desires and demand for goods and, second, a physical half of transporting and storing goods. The immaterial half creates "possession utility". The physical half creates "utilities of place and time" (Converse 1954). Alewell (1968) distinguishes the acquisition of demand and the physical distribution of goods. Correspondingly, Coughlan et al. (2006) distinguish the information and the fulfillment function of channels. As the activities of acquisition and physical distribution are inherently different by nature, the differentiation made by Converse, Alewell, and Coughlan et al. is helpful to understand distribution. Schögel (2012, 37ff) distinguishes the term channel: there is an acquisition channel for the immaterial half and a distribution channel for the physical half. These two can be integrated but do not necessarily have to[4]. Herein, the immaterial half of acquisition and information is interpreted as the focus of the marketing perspective. The physical fulfillment half of channels is the focus of the logistics perspective.

1.1.2 Logistics: the physical side of distribution

From the logistics perspective, the function of distribution is to bridge spatial, temporal, quantitative, and qualitative gaps in-between production and consumption (Ahlert 2005). Consequently, there are transfer processes to build these bridges: transportation, storage, commissioning (picking & packing), and sorting (Ihde 2001). The function of logistics systems is the coordination and the execution of those physical processes that are performed on a product to make it available.

Logistics as a managerial discipline and perspective emerged historically out of marketing (Converse 1954; Drucker 1962; Klaus 2002; Ahlert 2005). Marketing took a holistic view on distribution (Drucker 1962) and distribution-logistics has been identified as "the other half" (Converse 1954). One of the first to promote logistics as a driver of success and actually using the term was Magee (1960)[5].

Nowadays there are three definitions of logistics (Klaus 2002).

1. Logistics is the highly specialized optimization of singular "transfer" processes that add value of place, pace, and pattern to a product: the main transfer processes are the "three P's"

[3] Property rights are distinguishing usus, usus fructus, abusus, and ius abutendi. The term property right acknowledges that the value of a good does not only depend on the property, but on the rights and obligations that come with a good. The economic function of the term is the clear and transparent allocation of benefits and harms of goods and services (Demsetz 1967). For example, an angler possesses fishing equipment. In Germany, a state license is required to be legally allowed to use it for landing and killing fish. In order to allocate the harms of fishing (e. g. decrease in the fish population, damages on the riverside) on a particular water, he has to buy the permission from the holder of that right (often communities or unions) to go fishing on that particular water.

[4] For example, a showroom is a conventional outlet but does not carry any stock. It is solely used to acquire demand but not to fulfill orders. Instead, the ordered product is then delivered in the home delivery channel (Bell et al. 2018).

[5] Magee (1960) already outlined the most essential trade-offs in logistics management, that are discussed nowadays: product variety vs. inventories; distribution cost vs. response time; warehouse facilities vs. transport time and cost; make or buy transportation and warehousing functions; levelling employment, inventories, and capacities vs. demand variations.

transportation "place", warehousing "pace", and transshipment "pattern" (including commission, sorting, packaging etc.)[6].

2. Logistics is the managerial intra- and inter-company coordination of all activities that create utilities of the "seven rights": the coordination is exercised at the interfaces between production, transportation, transshipment, warehousing, and consumption.
3. Logistics is an organizational flow paradigm that interprets the supply chain as a system where elements perform their standard operating procedures on goods; thereby goods, information, cash, capacities, people, and ideas "flow" through the system. The flows are the objects of logistical decision-making.

The definitions developed over time, but do not necessarily build upon each other. They differ in their understanding of what logistics as a discipline does and what the object of logistics-management is. However, the function of making goods available is the same. Therefore, there is no more or less important definition, but different issues, methods, and goals. The third definition focuses on the chain-wide process-orientation, the second definition focuses on the interfaces between processes, and the first definition focuses on the optimization of singular processes.

The logistics perspective on distribution views distribution as one of three generic sequentially linked processes: procurement-logistics, production-logistics, and distribution-logistics. In Fig. 1.1, the interfaces with the supply and demand side are insinuated by the overlapping chevron boxes. Procurement-logistics manages the flow of all kinds of purchased inputs into the company. Inside the company, the production performs standard operating procedures transforming inputs into outputs. Production-logistics creates availability of the inputs throughout the procedure according to the production master plan. Distribution then takes over the outputs (finished goods) of production and manages the flows to the procurement of the demand-side. At this point it is clear, that the distribution of the manufacturer overlaps with the procurement of its customer and the procurement of the manufacturer equals the distribution of its supplier.

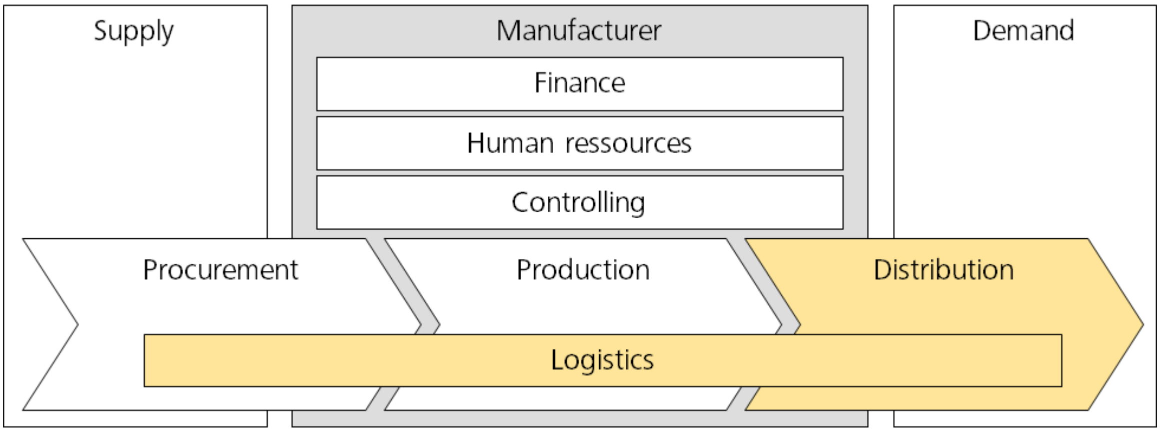


Fig. 1.1 Sequential processes of business logistics at a manufacturing company

[6] The three "P" of logistics "place", "pace", and "pattern" correspond to the German TUL="Transport, Umschlag, Lagerhaltung"-logistics.

Assuming the process-flow paradigm of logistics, distribution-logistics is defined as all activities that plan, enable, prepare, execute, manage, and control the flow of finished goods from a manufacturer through the channel towards its customers (Ihde 2001, p. 296; Specht and Fritz 2005, p. 115; Schögel 2012, p. 363). The task to configure these flows raises questions like:

- "How do products move through the channel?"
- "Where do products sit?"[7]
- "Who brings a product from the production site to the customer?"
- "What amounts of different products are bundled together?"

1.1.3 SCM: the chain "from farm to fork"

When SCM first emerged as a field of research, it was not clear at all, whether it is the same as logistics, or an extension, or an "all encompassing approach to business integration" (Cooper et al. 1997). Cooper et al. (1997) state that SCM is a broader concept than logistics management. However, it seems cumbersome to distinguish the third definition of logistics from Klaus (2002) and SCM. This discussion has no practical use or benefit.

One definition of SCM that has been sharpened over time stems from the Global Supply Chain Forum: "Supply Chain Management is the management of relationships in the network of organizations, from final customers through original suppliers, using key cross-functional business processes to create value for customers and other stakeholders" (Lambert 2014). The relevant insights of the SCM perspective stem from the chain-wide investigation and management of relationships, processes, flows of information, goods, and financing.

The new aspect is that the scope of discussion takes off from analyzing and optimizing intra-organizational processes, divisions of organization, decision-making, and their relationships. In SCM, the managerial focus lies on inter-organizational processes and their seamless integration. The scope of this integration is "from farm to fork", meaning from the point where something is farmed from the earth to the point where it is finally consumed. Another metaphor "from dirt to dirt" (Kumar 2001) includes further the processes of waste disposal and reverse logistics (recycling) in the scope of SCM. The chain-wide integration has two normative levels: first, the integration enables chain-wide visibility through information sharing, second, integration enables improvements (even optimization) of the chain-wide flows through cooperative planning.

The objective of visibility is more intelligent coordination across organizational boundaries. Corporate planning on every stage of chain is easier because visibility mitigates "nasty" surprises and detects early signs of trouble at these boundaries. As a result, uncertainty and risk get more and more replaced by robust promises, proactive prevention, and agreed "fire fighting" strategies. For example, geo fencing systems make the approach of delayed inbound transports visible. Knowing the actual time of arrival early enables adjustments in the receiving area: further delays through unavailable storemen can be mitigated. The second level of integration, cooperative planning, grounds on the insight that locally optimized decisions may lead to a global, chain-wide

[7] Inventories are "stuff sitting somewhere", cf. van Ryzin 2001, p. 1.

suboptimum. Therefore, the deployed resources (e. g. safety stocks, production and transportation capacities, shared information bases) and logistics processes (e. g. replenishment and ordering) should be optimized from a chain-wide perspective, instead of doing so at every stage locally (Lee et al. 1997). The ultimate objective of SCM is to minimize the sum of resources that are consumed and wasted along the supply chain "from farm to fork" in order to satisfy the demand (Houlihan 1985). This belief that integrated supply chains produce superior results is the paradigm that shapes the field of SCM[8].

SCM proposes intra- and inter-organizational "key cross-functional business processes"[9] which smooth the seamless flow of orders through the supply chain (Cooper et al. 1997). The property "cross-functionality" means functions collaborate across their borders in a way that creates more value-add than the vertical sequence of those functions could. Examples of such cross-functional business processes for the integration of the supply chain are Continuous Replenishment, Quick Response, Efficient Customer Response, Collaborative Planning, Forecasting and Replenishment. For more examples see Tyan and Wee (2003). All these processes have an information sharing aspect, which creates visibility, and a consumer-oriented aspect, which smoothes the flows through the chain: the chain "breathes with the customer". In practice, these processes are often implemented in tier 1 relationships, but they are aimed at a broader scope.

Looking at the supply chain as a whole, every firm is a sequence of the four generic processes procurement, production, distribution, and sales.

Fig. 1.2 embeds the supply chain planning matrix (Fleischmann et al. 2008, p. 87) in the inter-organizational flow-oriented view. Compared with Fig. 1.1, the "sales" process is an additional boundary-crossing process. Whereas Fig. 1.1 displayed the physical flows of goods, the supply chain planning matrix utilizes information flows to plan sales volumes and revenues. Fleischmann et al. (2008) disassemble marketing into distribution (physical) and sales, meaning the acquisition of demand (see sub-section 1.1.1). The matrix zooms into one focal company in the chain and observes the sub-processes along the two dimensions of time horizon and the process dimension. The customer order hits the supply chain at the Point-of-sale (POS). In a make-to-order environment (Otto 2008), it directly activates all processes of the supply chain in an upstream direction. In a

[8] Bretzke (2005) criticizes the chain-wide, holistic integration with the objective of optimization as utopia. He points out that there is a "price of supply chain management", namely, the renouncement of the spot market. Bretzke first observes that companies like METRO try to "lock-in" their single-sided market-power without sacrificing neither their buyer's rent nor their sourcing flexibility from multi-sourcing. In general, individual actors are unlikely to forgo opportunistic defection in "locked-in" relations. Second, he points out that the optimization of the whole supply chain is illusive, as it requires many mathematical relaxations, assumptions, and simplifications. The application of OR has been criticized for solving lower-level problems and thus sub-optimizing more complex decision problems* (Hitch 1957, p. 718; Lindblom 1959, p. 80). Third, since suppliers deliver to more than one customer and, vice versa, customers buy from more than one supplier, most companies are nodes in more than one supply chain. Therefore, a hypothetical global optimum of one supply chain is interdependent with many other supply chains that intersect in this node: it is really a supply network. For example, a producer of carbon fibers has customers in the automotive, the security textile, and the fishing sports supply chain. It is thus a node in all three supply chains, which become interdependent through this connection.
* A pragmatic would likely agree and argue that it is nonetheless better to regardlessly apply simplified optimization models than none (see sub-section 1.4.1). By the terminus "pragmatic" I mean somebody who is willing to sacrifice some aspects of due diligence on the altar of "getting things done". That kind of mindset may be studied in Lindblom 1959 on "muddling through complex problems" and in Kirsch 1998, pp. 76–77 on "Das Durchwursteln".
[9] Herein, a business process is understood as in Cooper et al. 1997, p. 5: „A process is a specific ordering of work activities across time and place, with a beginning, an end, and clearly identified inputs and outputs, a structure for action. Supply chain business processes can cross intra- and inter-organizational boundaries, independently of formal structure."

make-to-stock environment, the order is served from finished goods inventory and indirectly triggers the replenishment of stock. It wanders from the procurement processes of the downstream company to the distribution processes of the upstream supplier and from the supplier's procurement processes to its next tier supplier's distribution processes and so on (Otto 2002, pp. 93–94). In a perfectly integrated world, the farmer is informed in real-time about what the consumer has "on the fork".

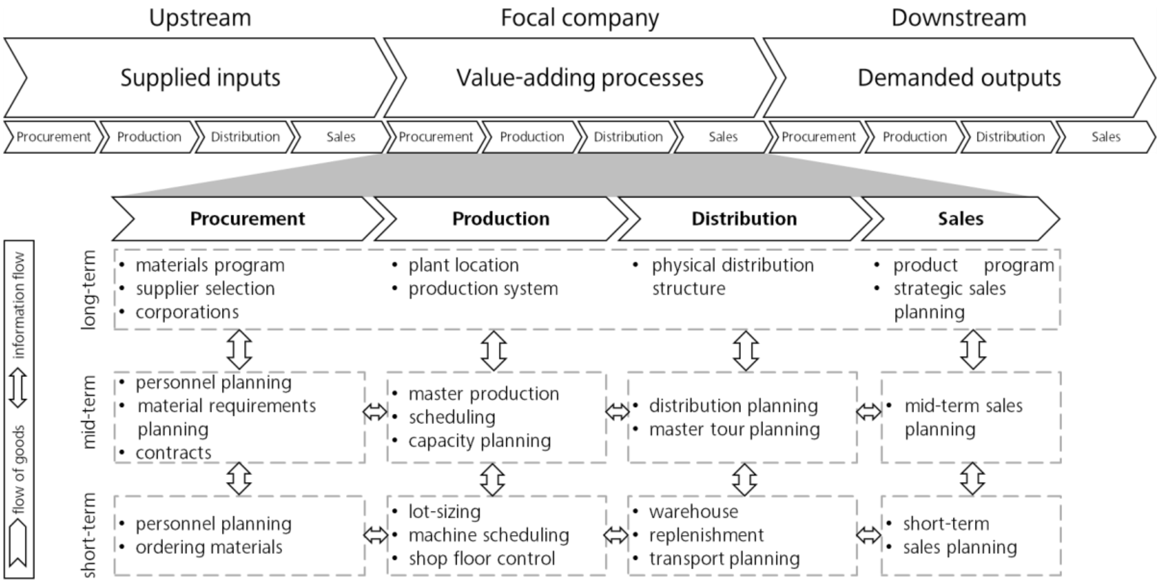


Fig. 1.2 The supply chain planning matrix in inter-organizational processes

How does SCM view distribution? SCM views distribution as a subsystem of the supply chain system that performs all processes that relate a company's production with its customers' procurement processes. From an SCM perspective, distribution has the function to integrate the flow of finished goods from productive processes towards consumptive processes. In other words, distribution transfers unavailable goods into available goods. Distribution has the planning processes itemized in Fig. 1.2 classified by their time horizon. The structure of this sub-system consists of warehouses and other nodes and the transportation links in-between them (Fleischmann et al. 2008). The scope of SCM when discussing distribution is not only on the motion of goods "from farm to fork" alone, but on the engineering of information flows through the stages of the supply chain in a way that the customer's order can wander seamlessly in the upstream direction "from fork to farm".

1.2 Distribution-logistics for consumer goods

This section describes the distribution of consumer goods from the logistics perspective. Herein, the chain-wide flow ("How do consumer goods flow from the upstream sources down into the customers' hands?") is the broader paradigm of the investigation. The later sections of this dissertation have a much more narrow focus on the transportation on the very last stage of the distribution, the so-called "last mile".

1.2.1 Research focus: distribution of consumer goods

The distribution processes differ according to categories of distribution objects. Goods are categorized by various criteria[10]. Typical categorizations are input vs. output goods, industrial vs. consumer goods, consumption vs. durable goods, parts vs. finished goods, physical vs. immaterial goods, and real vs. nominal goods (Thommen and Achleitner 2006, p. 35). This dissertation deals with and limits itself to physical consumer goods, as opposed to industrial goods. Therefore, raw and semi-finished materials for production, services, digital products, passengers, etc. are not discussed. The categorization of industrial and consumer goods distinguishes if a product satisfies a final consumer demand directly or indirectly. Consumer goods satisfy final consumer demand directly, industrial goods are inputs (including investment goods and materials) into industrial operations and thus satisfy demand indirectly (Thommen and Achleitner 2006, pp. 35–36).

What are consumer goods (CG)? CG are used and consumed by the final customer. They are often characterized by high standardization, low functional complexity, and modest difficulty for handling. CG are further sub-classified as fast moving (FMCG) and slow moving consumer goods (SMCG). On the one hand, FMCG are products of daily use, e. g. groceries and sanitary products that are bought and consumed high-frequently, are easily substitutable and non-durable, have a basic price and narrow profit margin, low customer engagement, low risk when purchased, and short shelf-life duration (Ahlert 2005, p. 44; Schögel 2012, p. 374; Otto et al. 2018, p. 738). On the other hand, SMCG are durable products of sporadic, non-frequent purchase, e. g. white goods, consumer electronics, and musical instruments, that impose a high risk of mispurchase to the customer due to higher technical complexity, are hard to substitute equivalently, have a long shelf-life and thus impose risk of depreciation to the inventory holder (Ahlert 2005, p. 44).

1.2.2 On the economic relevance of transportation

In Germany, CG are usually distributed on the road, as opposed by rail, water, or air. This sub-section highlights the predominant role of transportation on the road in the overall German logistics market. Focusing on road transportation, goods are transported in four different modes that vary in their typical shipment size. In order of decreasing shipment size, these are: full-truck-load (FTL), less-than-truck-load (LTL), groupage freight (GF), and courier, express, and parcel (CEP). The smaller the shipments are, the more shipments are bundled together in transportation in order to exploit economies of scale in transportation, herein called "economies of transportation" (see sub-section 1.2.4).

[10] Generally, objects of distribution can be classified by several of their properties: physical attributes (weight, volume, perishability, aggregate state), technical properties and individualization, complexity, value-added service bundles, price level, frequency or pace of sales, urgency or priority, phase of product life cycle, digitalization of the product Specht and Fritz 2005, pp. 219–220.

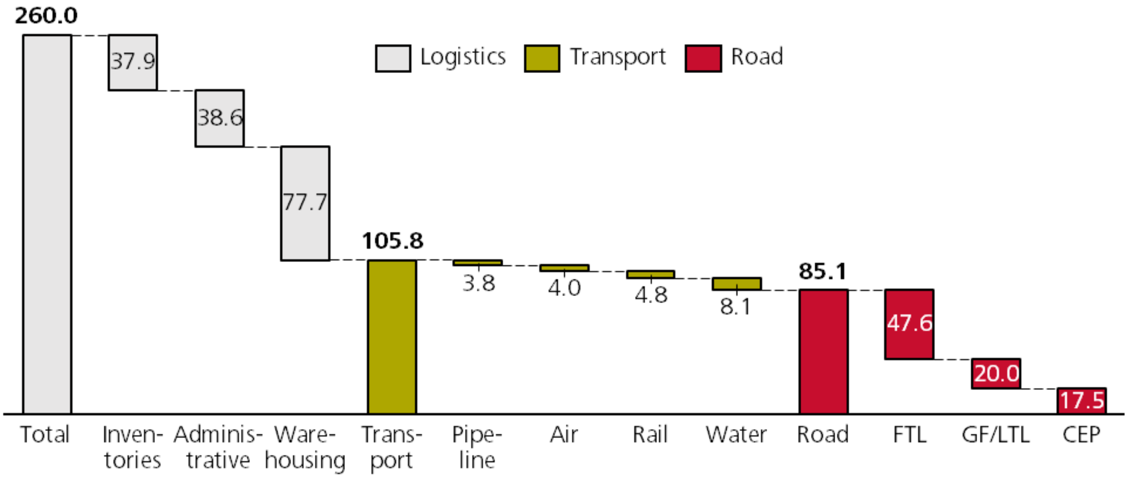


Fig. 1.3 Cost volume of logistics in Germany, 2015, in bn. EUR, (numbers selected from Schwemmer 2016; visualized by the author)

Fig. 1.3 displays estimates of the total volume of logistics costs in Germany in 2015 (Schwemmer 2016, 60–71, 83-87). The total cost volume has been EUR 260 bn. Transport is the largest fraction having EUR 105.8 bn. (40.7%) of the total. Adding the direct administrative cost of transportation (the total admin. costs have been EUR 38.6 bn.), the fraction of transportation increases to EUR 116.8 bn. The major share of transport is "on the road" having a total volume of EUR 85.1 bn. (80.5%). The transportation market is further segmented into thirteen submarkets. In Fig. 1.3, the relevant submarkets "on the road" are clustered by typical shipment size. Full-truck-load (FTL) consists of truckload (EUR 27.1 bn.), heavy duty (EUR1.2 bn.), tank & silo (EUR 7.8 bn.), and special equipment (EUR 11.5 bn.). Groupage freight[11] (GF) having EUR 11 bn. and dedicated groupage freight networks having EUR 9 bn. are pooled under groupage freight and less-than-truck-load (GF/LTL). CEP logistics is one of the thirteen submarkets accounting for EUR 17.5bn.

1.2.3 Strategies for distribution networks of consumer goods

The major characteristics of the distribution-logistics of CG are (a) the deployment of intermediaries and (b) few-to-many divergence on the later stages of the distribution chain.

CG are usually distributed indirectly, this means via intermediaries like wholesalers and retailers. The role of intermediaries is to bundle products from many sources in their retailing channels and add additional services to consumer goods (Brown et al. 2000). Bundling enables degression of distribution costs per unit. Further, the value-added services that are performed in an outlet require a critical mass in order to reduce process costs per unit. This critical mass is achieved through bundling. For example, if all the CG that are sold in a discounter outlet were sold in separate channels, then the physical structure and the value-added services would be duplicated. Therefore, most CG manufacturers sell through intermediaries[12].

---

[11] Schwemmer 2016 equates groupage freight with less than truckload.
[12] Original manufacturers of high-end SMCG like Vorwerk or Apple run own outlets. Other comparable manufacturers like Stihl go a middle way by subcontracting with exclusive specialist retailers. The rational is that the products are profitable

On the last mile of distribution, products are distributed into the area in relatively small order volumes to many small and often anonymous customers. The so-called last mile problem is how to get a product "into the consumers' hands" (Larke et al. 2018, p. 466). The last mile and the issues associated with it are further elaborated in subsection 1.3.

How are CG physically moved? Very large FTL shipments are transported directly to a retailer's warehouse. Direct shipping does not require own networks, since FTL shipments are not transshipped. Smaller LTL shipments and even smaller shipments that require consolidation in groupage freight are transported in networks. Logistics service providers (LSP) operate networks that are often specialized on transporting a certain type of shipments. Networks of LSP differ in the number of terminals and hubs. In Germany, the number of terminals is typically 23-35 (frozen goods), 26-28 (fresh goods), 45-50 (dry goods) (Otto et al. 2018, p. 741). The shipments move through the LSP's network towards the customers. From a manufacturer's perspective, the customers are wholesalers and retailers[13]. The shipments are delivered to those customers' own logistics networks consisting of warehouses, cross docks, outlets, and other types of sales points. From the retailers' warehouses, the shipments are further distributed through the channels towards the final customer.

In practice, blue-print strategies of CG distribution have emerged. Otto et al. (2018) elaborate three distribution strategies focusing on FMCG: central warehousing, direct store delivery, and cross docking (Fig. 1.4). The typical distribution strategy for SMCG is central warehousing. The key difference between slow and fast moving CG with respect to distribution is the tendency to centralize inventories.

Central warehousing is characterized by the manufacturer and retailer each operating a dedicated network of warehouses. This strategy centralizes CG inventories in very few central warehouses (CWH), from where they are further distributed. Both conceptional analyses and practical experience show that manufacturers of SMCG operate only one warehouse for Germany or even for Europe. Differently, manufacturers of FMCG usually operate more, typically four to five, regional warehouses (RWH) in Germany (Kellner et al. 2013). For example, Kellner et al. (2013) report that a large German manufacturer of dry FMCG runs three warehouses from where the German retailers are supplied. Shipments of CG are then handed over to the retailers' network of warehouses. These shipments are usually LTL and consist of multiple homogeneous pallets as opposed to mixed pallets. The transfer in-between retailer and manufacturer is performed either by the manufacturer or by the retailer or by a subcontracted LSP. The retailer distributes the goods further through the retailing channels into the final customers' hands. That sort of multi-echelon warehousing is conventionally deployed, because manufacturers usually supply many retailers and retailers usually procure from many manufacturers. The role of CWH is to consolidate those flows of goods (Daganzo 1987).

enough to bear the additional cost, but on the other side customers value the additional (after) sales services like advisory, financing, packaging, repairing, or cross selling.

[13] Otto et al. 2018 provide some figures from publically available studies about the different types of retailers.

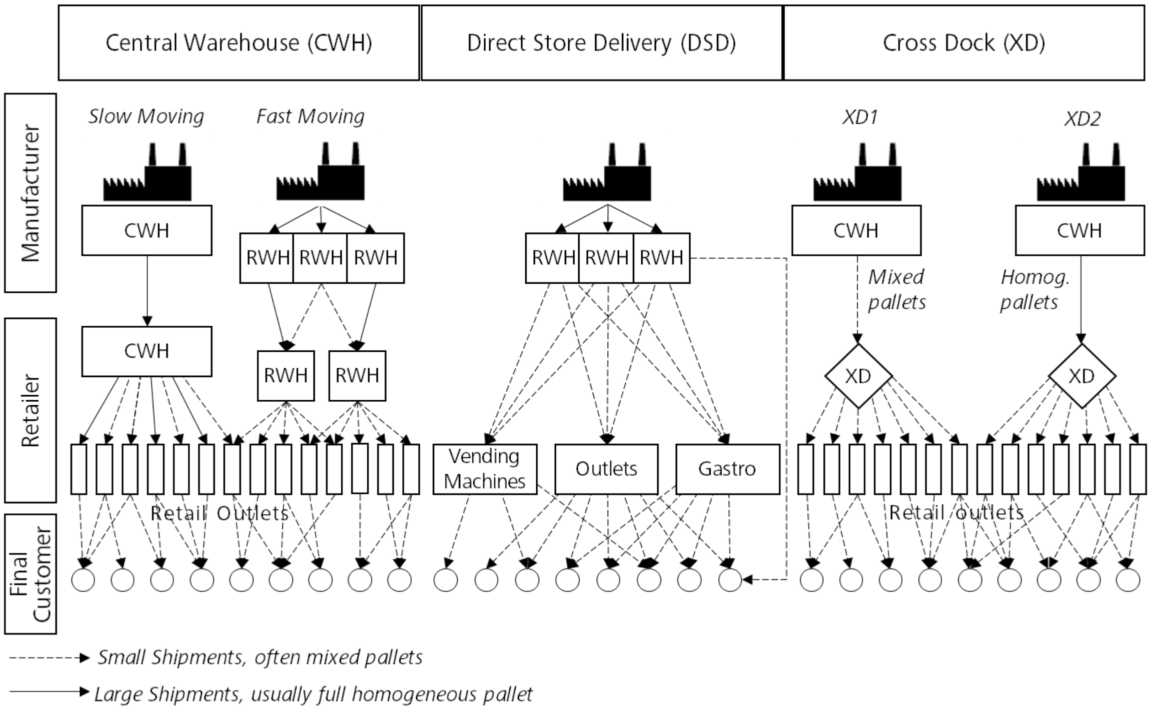


Fig. 1.4 Generic strategies of distribution-logistics for consumer goods (Otto et al. 2018; translated by the author)

Direct Store Delivery (DSD) is a strategy that forgoes the retailer's warehouse and thus also forgoes the consolidation role of that echelon. The manufacturer takes over the responsibility to supply the outlets with mixed pallets instead of homogeneous pallets. The distribution tour starts at the manufacturer's RWH and approaches many stops. These stops are many retail stores or gastro outlets or vending machines, instead of only few regional retail warehouses. As a result, the size of delivered shipments per stop is way smaller than in central warehousing.

The rational to forgo the consolidation at the retailing warehouse and employ DSD anyway are the reduction of distribution echelons leading to improved replenishment cycle times from manufacturer to shelf, the customer exposure (e. g. taking orders when delivering, communicating with the customer), market research (e. g. notice which competitors are on the shelf, too), providing value-added services (e. g. building up displays of merchandise), and ensuring delivery service quality (e. g. having own qualified staff instead of an LSP) (Otto et al. 2018).

A **Cross dock** (XD) is a retailer's regional transshipment point without any storage capabilities. In a XD, there is only the pattern of a shipment transferred (commissioning), but not its pace (no temporal transfer through warehousing). There are two process variants. The difference between XD1 and XD2 is who is responsible for the commissioning process for the retail outlet: in the case of XD1, the manufacturer owns the commissioning process in its CWH; in the case of XD2, the retailer owns it in the XD (Otto et al. 2018). In practice, also hybrid forms of central warehousing and cross-docking are employed, where there is a cross-docking area within a CWH. Thereby, some items can bypass the storage process, while others are sitting there waiting to be forwarded.

1.2.4 Economies of Transportation: the essential trade-off between bundling and sorting

Economies of scale in transportation follow special mechanics that are herein called "economies of transportation". Understanding this concept and its underlying trade-off is key to configure a distribution system with the objective of efficient distribution. The central trade-off is between bundling (elsewhere "pooling", e. g. Rouquet and Vauché 2015) and sorting. On the one hand, bundling is an LSPs central tool to reduce the transportation cost per unit. On the other hand, bundling requires sorting or consolidation processes that consume time and lockup resources. Economies of transportation guide decision-makers in their mission to make transportation more efficient, less costly, create more value from transportation.

The transportation costs of a single shipment increase by duration, mileage, and the shipment's size[14]. The marginal cost per kg are decreasing due to cost degression. But why do not all shipments have an enormous size then? Because the place, time, and amount of demand is spread. Stock-keeping of very large shipments is associated with costs and capital expenditure; too large stocks congest warehouses, and stocks lock up cash and thereby diminish liquidity. Someone has to bear those inventory costs. Logistics trades off the costs of transportation and the costs of inventories (Tempelmeier 2015). Even if a manufacturer wants to move FTL because it is cheaper, the customer is not willing to bear the higher inventory cost.

From an LSP's perspective, there is usually no shipment-sizing decision to make – the size is predetermined by the shipment order and the transportation service is a reflex to that order. The sizing decision is a privilege of manufacturers and retailers, but not of LSPs. Making the trade-off between large shipment sizes and small inventory lots, manufacturers and retailers use OR methodology to find an optimal lot size and thereby optimize their total cost (Andler 1929; Reichwald and Dietel 1991; Chopra and Meindl 2014). Instead, LSPs try to bundle multiple shipments in order to leverage economies of transportation. There are three typical ways to do so: bundling of ship-from addresses in an origin area, bundling of recipient ship-to addresses in a delivery area, bundling of freights on the line-haul from the same shipping terminal to the same receiving terminal or hub (Fig. 1.5).

[14] The size of a shipment may be a measure of weight, payload, volume, "tax weight", or number of items. Unless otherwise stated, I mean the weight of a shipment.

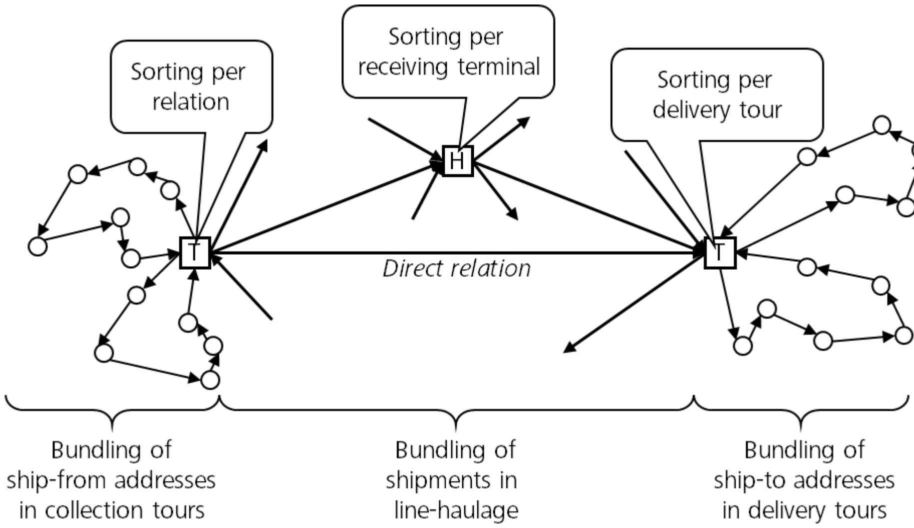


Fig. 1.5 Typical patterns of bundling in an LSP network (Bretzke 2008; modified and translated by the author)

However, bundling comes at some costs. An LSP bundling ship-from addresses needs to perform a collection tour (pre-carriage). An LSP bundling ship-to addresses needs to perform a delivery tour (onward-carriage)[15]. In GF and CEP networks both collection and delivery tours are necessary (Bretzke 2008). Furthermore, bundling requires sorting processes in every terminal and in the network's hub. Sorting is especially time consuming because sorting has to wait for the arrival of the last inbound relation to hand over its shipments. These additional sub-processes – pre-carriage, onward-carriage, and sorting – make the transportation of small shipments much costlier. Fig. 1.5 systemizes those additional processes. The configuration of distribution systems exploits bundling as much as possible to the extent when these additional costs exceed the benefits.

1.3 The last mile of distribution-logistics

The term "last mile" has different, context-specific meanings. It originates from the time when private households first got access to the nearest local exchange. In telephone parlance, the part of the telephone network dedicated to a single connection was called "the last mile". Nowadays "last mile" diffuses from wired telecommunication into the parlance of almost every network: distribution, humanitarian relief, public transportation, power- and gas supply, wireless communications, health and medical care. Also sports medicine and military operations use the term "last mile" metaphorically as the hardest mile. This chapter intends to explore an understanding of the last mile in distribution-logistics of consumer goods; what the "last mile problem" is; and why the last mile is considered the hardest mile.

[15] Research on freight forwarding road transportation is driven by contributions in German language. As a result, translations of freight forwarding parlance into English are ambiguous. Pre-carriage means „Vorlauf", "line-hauling" means "Hauptlauf", and onward-carriage means „Nachlauf" in German.

1.3.1 The meaning and the scope of last mile distribution-logistics

Defining what the last mile in distribution-logistics actually is, is difficult without a specific application case. Some authors refer to the last mile without defining what it explicitly means, taking a common understanding for granted (e. g. Punakivi et al. 2001). Most authors refer to the last mile within a specific context, usually the fulfillment of e-commerce orders through home delivery (e. g. Lee and Whang 2001, Chen and Pan 2016, Macioszek 2018).

1.3.1.1 Examples

It seems difficult to pin down a general definition of what the last mile and the so-called last mile problem are. In practice, many different cases and settings from distribution can be interpreted as last mile distribution. The following examples give an insight into the scope of different cases in last mile distribution.

1. Last mile as the local distribution of CG using a private fleet: the production site of a regional bakery chain, e. g. Ebner, produces fresh bread loafs on a daily basis. The loafs are commissioned from “the dry end of production” into standardized transport baskets (Euro measure 600 x 400 cm). Small CoDi[16] vehicles distribute the baskets over the last mile to many small dispersed bakery outlets across the service area.

2. The last mile as gastro meal delivery using own couriers: small, local gastro shops like pizzerias offer a local home delivery service of their meals. The meal is packed in cardboard boxes and shipped by employed couriers using cars, scooters, or bikes. The service is free of charge but there is a minimal shipment value and maximal distance enforced.

3. Last mile as the last formal stage of the LTL/GF process: a freight-forwarding company, e. g. Streit+Co, is a member of a groupage freight network (cooperation), e.g. CargoLine. Every member performs pre- and onward-carriage in its service area. All shipments into the company’s service area are line-hauled from the network’s central hub to the firm’s terminal. This terminal serves as the central facility in the service area and large rigid CoDi[17] vehicles cover the last mile and deliver the shipments to their final destination.

4. Last mile as city logistics: a city LSP runs a hub at the gates of a city. The basic idea of city logistics is that many freight carriers who forward shipments into the city hand over all of their shipments to the city LSP. This city LSP can achieve a higher drop-factor and thus better economies of transportation through bundling of ship-to addresses in the city. From the hub the city LSP performs the onward-carriage and thus covers the last mile to all stops in the city using a smaller number of CoDi vehicles than the total of all freight carriers.

5. Last mile as in-house delivery of mail: the internal postal service of a large firm or authority, e. g. university, takes ownership of the in-house last mile distribution process. A CEP/mail service provider, e. g. Deutsche Post, hands over a large number of bundled letters and parcels at the

[16] CoDi=“Collection and Delivery”. A small rigid CoDi has a typical payload of 3.5 tons.
[17] A larger CoDi has a payload between 5.5 tons (“12-Tonner”) and 10.0 tons (“18-Tonner”).

university's post office. The in-house distribution to employee offices takes a lot of time because offices are unattended, delivery is done by foot, or opening hours of secretaries differ vastly. Therefore, the post office is also a pick-up point.

6. Last mile as the last stage of parcel home delivery: Consumers order products from retailers or manufacturers who pack the customer order and ship it with an LSP. The LSP, e. g. DHL, Hermes, or UPS, collects several shipments from every shipper and consolidates all shipments over several transportation legs toward their general directions. In a regional terminal the shipments are re-consolidated into a delivery tour that approaches every single destination within a regional area.

7. Last mile as the replenishment of retail outlets: a large German discounter, e. g. Lidl, consolidates FMCG from many manufacturers in its RWH. In the RWH, the picking and packing of mixed pallets for the supermarket outlets is highly automatized. The retail outlets are replenished daily with multiple mixed pallets. Lidl runs its own fleet of large CoDi trucks and tractor-trailers in the replenishment process.

8. Last mile as the fulfillment of grocery e-commerce: a large German supermarket, e. g. Rewe, operates an e-commerce channel that allows customers to order their groceries online. The ordered items are picked and packed in a regional fulfillment center and delivered either to the customer's home or to an outlet, where the order waits for the pick-up.

These eight examples show the key commonality and characteristic of last mile distribution cases: the one-to-many transportation process. This corresponds to the right-hand side of Fig. 1.5. There is "one" break-bulk point (BBP)[18] that can be identified in all examples: it is the point where hitherto bundled flows of goods are re-consolidated into delivery tours. The BBP in the examples are the central bakery, the pizzeria, the receiving LTL terminal, the city hub, the mail receiving office, the regional terminal, the retailer's RWH, and the fulfillment center. The BBP is herein defined as follows:

> The BBP is the point in the distribution process which divides the distribution in a part, where all shipments having the same general direction are handled together in bulks, and a "last mile" part, where all shipments are re-consolidated by their distinct destination and delivered individually.

The "to-many" are the various points of delivery: touchpoints with the customer, where the physical hand-over takes place. Koether (2014, p. 39) states that the last mile distribution from the BBP to touchpoints with the final customers is always the same process for different distribution systems[19]. Nevertheless, there are differences in the examples and thus the last mile process – inherently the same – varies with respect to the specific context.

[18] I understand the terms "break-bulk point" and "break-bulk terminal" (e. g. Daganzo 1987), synonymously. Herein, the term "break-bulk point" (*German: "Vereinzelungspunkt", "Dekonsolidierungsknoten"*) is used.
[19] Couriers are an exception from the rule. Couriers offer a guaranteed door-to-door service without any transshipments and switches of the responsible carrier (Bretzke 2008). Therefore, courier services lack the BBP and to not distinguish a line-haul and a last mile.

1.3.1.2 Systematization

In order to develop a definition of the last mile, the different cases that are addressed as last mile problems are hereafter systematized. The dimensions are (a) who the responsible actor is and (b) where the point of delivery is.

Dimension (a): Who is responsible for the last mile?

Last mile distribution is covered by either the manufacturer, who makes the product, or by a retailer, who sells the product to the customer, or by an LSP. In case of a responsible retailer or manufacturer, an LSP may be subcontracted in order to take over the operations and few coordinating and managing activities. The question of interest is, who acts as a planner, designer, coordinator, and integrator on the last mile?

Dimension (b): Where is the point of delivery?

The point of delivery means that point in the process of making a good available, at which this good is physically handed over into the hands of the customer. Different points of delivery have emerged on the last mile. In general, there are conventional "brick & mortar" outlets, where CG are picked from the shelf, and there is home delivery to the customer's doorstep. Recently, e-commerce enabled the development of a spectrum of "in-between" strategies leading to pick-up points as a third point of delivery. The three strategies differ especially by their customer involvement in the distribution process. Customer involvement means the scope of activities that customers are expected to do themselves.

(1) Conventional "brick & mortar": this pick-up strategy is applied for most FMCG and ends in the retail "brick & mortar" outlet. The consumers perform many logistics activities themselves in the shopping cart: selection (right product), commission (right amount), packaging (right quality[20]), and transportation (right place). Customers bear the cost of transferring a CG from the outlet to the point of final consumption. The customer involvement is thus relatively high and the touchpoint with the final customer is closer to the production (Schögel 2012; Otto et al. 2018; Kuhn et al. 2018).

(2) Home-delivery: mail-ordering and e-commerce decouple ordering and fulfillment. Conventionally, consumers have to be physically present in an outlet to declare their intent to buy. As this necessity fades away, customers can make their purchase everywhere and the product is shipped to the customer's home (attended and unattended). Therefore, customer involvement in the physical fulfillment process is minimal and the service level provided is high in this ship-to strategy.

(3) Pick-up in-Between: e-commerce enables new strategies, how to bring the product into the customers hands[21]. "Click&Collect" or "Click&Reserve" are other common names for these

[20] In a broader sense, packaging redounds on the quality of products through protection, temperature control, and better handling.

[21] The development of in-between strategies that are enabled by e-commerce are driven by the "Amazonization" (Hotz and Fost 2016; Stüber et al. 2018) of customers. This coinage means that Amazon shapes the strategies for the last mile progressively. The "in-between" pick-up solutions are fine-tuned for customers' service demands. The last mile delivery is

strategies (Kuhn et al. 2018). Shipments are delivered to pick-up points. A pick-up point may be e. g. convenient stores (e. g. 7.Eleven, Larke et al. (2018)), public parcel lockers (e. g. DHL, Vakulenko et al. (2018)), or some depot box at the customer’s home or workplace (e. g. Night Star Express, section 3). The use of pick-up points changes the customer involvement in last mile distribution (Vakulenko et al. 2018). At the extreme, outlets or boxes at the customers’ homes can serve as pick-up points. Customer involvement and service level is gradually in-between conventional distribution and home-delivery.

Fig. 1.6 insinuates these three last mile strategies. The point of delivery (outlet, pick-up point, or home) is located somewhere between the BBP and the customers’ homes. The relation between distance from the point of delivery to the customer’s home and the customer involvement is highlighted. The closer the last mile reaches towards the customers’ homes, the more stops are on the last mile. At the same time, the distribution tours take more time for both driving and stopping. Consequently, these tours have higher costs per stop as a percentage of the delivered shipment value or the LSP’s revenue. The issue of last mile costs will be further examined in sub-section 1.3.2.1.

Applying the two dimensions (Who and Where to) of last mile distribution, the following Table 1.1 provides a systematization of cases that can be seen as last mile situations. The firms that are named in the matrix are picked as examples for the sake of the viewer’s orientation, but do not have any further relevance herein. The eight examples provided in the last sub-section are located in the table by their numbering. The ratings on the bottom and the right-hand side are estimations by the author and read + “high” and – “low” and -/+ “may differ”.

extended by innovative options that overcome the problem of unattended delivery. Amazon experiments with its own lockers and a door-lock that enables deliverymen to enter a customer’s private space, like the living room, garden shed, or trunk. Geo-tracking devices enable a “receiptless delivery” in B2C delivery (see the case of in-night express in section 3). Amazon experiments with its own fleet, drone delivery, bikes, taxis, and other unconventional LSPs for last mile transportation (Stüber et al. 2018). Mobile customers can receive their delivery in a more flexible way in terms of time and place. The carrier selection is increasingly adapted to different customer demands, attendance, time windows, mobile locations, and retour channel options. Consequently, the customer experiences greater power to decide how the last mile is bridged. Overall, the variety of delivery and retour options on the last mile explodes.

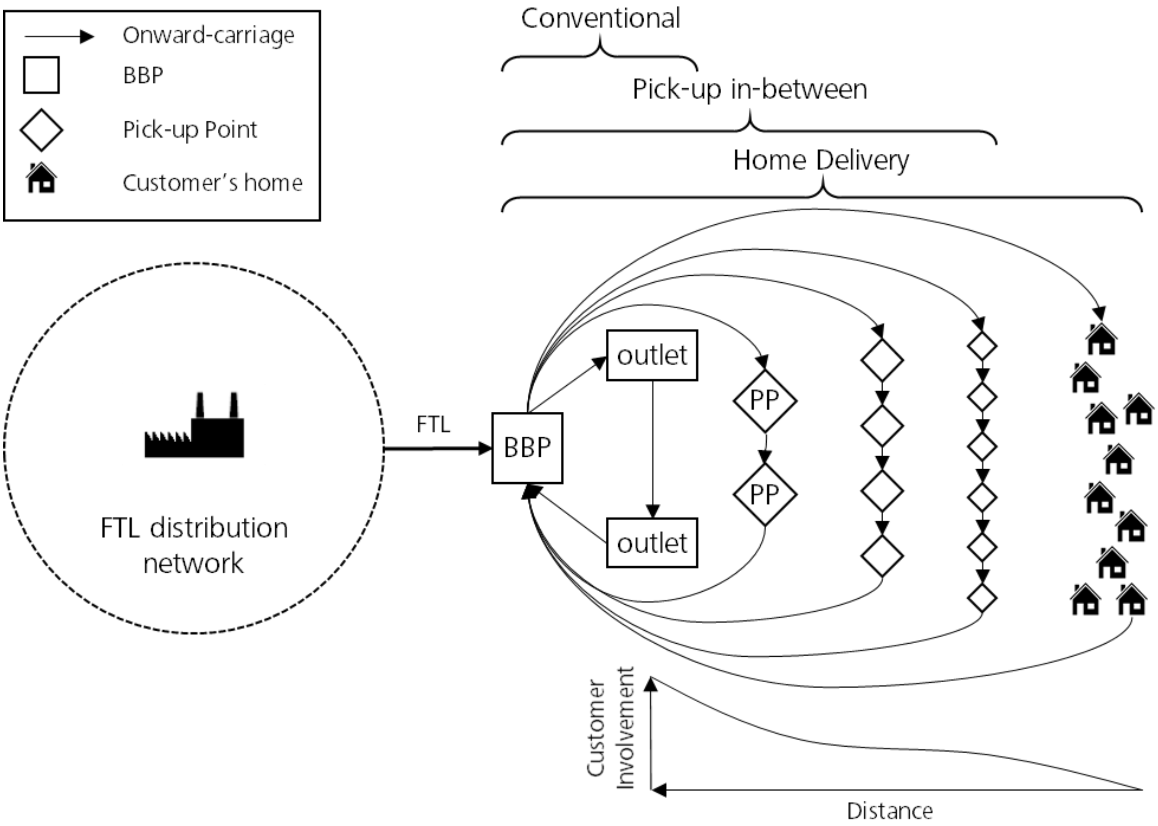


Fig. 1.6 Types of points of delivery on the last mile

The potential to decrease the costs per shipment through bundling is a general estimation of how great the bundling potentials of the different actors are. It is no assessment of the individual capabilities to actually identify and exploit given potentials. Manufacturers do not have much potential to bundle: only shipments of geographically close customers can be consolidated on a tour-level. Large manufacturers (e. g. Coca Cola) can bundle multiple products in their direct store delivery process. Another potential stems from cooperation among a club of manufacturers who supply the same set of retailers ("club pooling", Rouquet and Vauché 2015). If manufacturers decide to organize the distribution-logistics themselves, then they may subcontract LSPs for the last mile operations. In this case, the role of the LSP is an operator and mediator among the club of shippers.

Retailers have more bundling potential. Large retailers can consolidate products from many suppliers on a big-box-level (e. g. Carrefour does "domination pooling", Rouquet and Vauché 2015). In general, retailers can bundle small stops on a tour-level (e. g. Stahlgruber). In addition, retailers can over-deliver "smart size" shipments to utilize idle capacities and thus consolidate even more units[22]. LSPs have the largest potential to bundle on the last mile. They can bundle shipments for the same destination on a stop-level and for geographically close destinations on a tour-level ("customs pooling", Rouquet and Vauché 2015). If a retailer subcontracts an LSP for the last mile, the LSP has

[22] Smart sizing means to use empty transport capacities by over-delivering selected product groups. One rational of smart sizing is to over-deliver products, which are forecasted to be re-ordered in the short-term anyway and these future demands are delivered to outlets speculatively. For example, Jafari et al. (2016, p. 454) report that the discounter Lidl distributes non-food campaign items to the stores speculatively together with other regular CG, in order to use idle truck capacities.

the role of an operator and sometimes may manage some segment like a region on behalf of the retailer.

Small and medium-size manufacturers and retailers do not have the volumes per stop and thus struggle to distribute their products efficiently themselves. If the value of delivered CG per stop is small (e. g. one handling unit), then costs per stop are surely too expensive (Müller and Klaus 2009, p. 157). Therefore, that kind of manufacturer hands over the responsibility to an LSP who bundles stops in the collection and on the last mile on a tour-level ("district pooling", Rouquet and Vauché 2015). The number of shippers, shipments, and recipients is very large in an LSP network. Due to these large numbers, the consolidation and densification level in LSP tours and stops is much higher than the level of retailers and manufacturers.

The columns on the right are estimations of the number of stops per tour, the costs per stop as a percentage of the stop value or the LSP's revenue, and the involvement of the final customers. From top to bottom: the larger a stop is, the larger is the typical volume and the value per stop, the lesser stops are there. As a result, the greater the stop value is, the greater is the revenue of the LSP. This leads to a degression effect of the cost per stop. This effect is demonstrated and discussed in subsection 1.3.3. The greater the distance between the point of delivery and the customers' homes, the greater is the final customer involvement (Fig. 1.6).

Table 1.1 Systematization of types of last mile distribution

Who?

Where to?		Manufacturer	LSP	Retailer	Number of stops per delivery tour	Cost per stop as a % of stop value / of LSP revenue	Final customer involvement
B2B	Big Box Outlet	*Coca Cola*[23]	3 *KV Nagel*	*Carrefour*[24]	- - -	- -	+ + +
B2B	Small enterprises [25]	1 *Bakery Ebner*	4 *CargoLine NSE*	7 *Lidl Stahlgruber*	-/+	-/+	+ +
B2C	Pick-Up Point		5 *DHL locker box Qool Collect*	*S&I/ 7Eleven*	-/+	-/+	+
B2C	Home Delivery	2 *Local pizzeria delivery*	6 *DerliveryHero Post, Hermes*	8 *Rewe.de Amazon Fresh*	+ + +	+ + +	- -
Potential to decrease the cost per shipment through bundling		-	+ +	+			
Typical role of LSP (if involved)		Operator, mediator	Network planner, integrator, operator	Operator, managing on behalf			

\# Example of last mile problem from sub-section 1.3.1.1

[23] Müller and Klaus 2009 report of Coca Cola as an example of a CG manufacturer who pioneered DSD.
[24] Rouquet and Vauché 2015 report of Carrefour running collaborative consolidation centers, where the vendors are required to consolidate their stocks.
[25] Small businesses comprise all kinds of small local enterprises serving final customers, for example pharmacies, auto repair shops, mom-and-pop stores, gas stations, or bakery outlets.

1.3.1.3 Definition

This sub-section is an approach to carve out the essentials of a general definition of the last mile. The last mile addresses the final part of the physical distribution network, respective the supply chain. Hitherto bundled shipments are decoupled and re-consolidated to delivery tours. This decoupling and re-consolidation point is called the break-bulk point (BBP). The BBP divides the "few-to-few" line-haul and the "one-to-many" last mile part. Typically, the BBP is supplied by efficiently bundled "bulks" of shipments on the inbound side. These bulk flows are handled together by their general direction. In the BBP, these bulks are broken into singular shipments and re-consolidated in delivery tours (Daganzo 1987). The assignment of a single shipment to a distinct tour depends on the shipment's distinct destination. As a result, loads of delivery tours are composed of many small shipments and addressed to many small destinations (Punakivi et al. 2001; Edwards et al. 2010). On the outbound side of the BBP arises the full complexity of the significant level of dispersal of destinations. Different types of destinations are collectively called points of delivery, comprising "brick & mortar" outlets, pick-up points, and customers' homes. The last mile of the distribution of consumer goods[26] is then defined as:

> The last mile is the institutional, geographical, and temporal distance between the BBP and the many points of delivery in one-to-many distribution situations.

The so-called last mile problem is how to distribute CG from the one BBP to those many small points of delivery (one-to-many distribution problem). The above definition of the last mile highlights the beginning, the BBP, and the end, the point of delivery, of the last mile. In-between these points, there is a one-to-many transportation process that bridges the geographical distance. The temporal distance is bridged at the BBP by the sub-processes interim storage, picking, packing, and loading. The institutional distance is bridged at the start of the last mile during loading (hand-over to a subcontractor) or at the end (hand-over to the final customer). The hand-over process consists of all tasks that are performed between the vehicles handbrake is applied and released. Accordingly, last mile consolidation, transportation, and hand-over should be investigated as separate sub-problems of the last mile problem (Fig. 1.7):

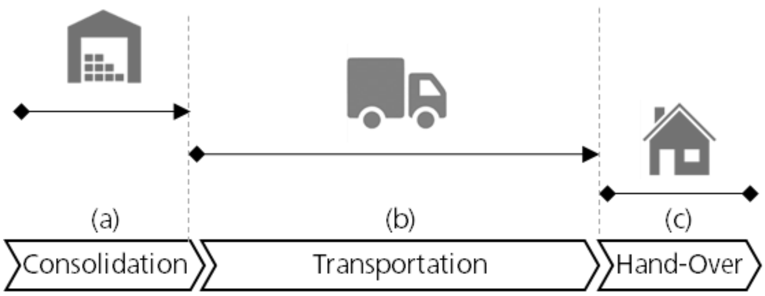


Fig. 1.7 Sub-Problems of the last mile problem

[26] A similar definition can be given for industrial goods, digital goods, real estate, property rights, services, and passenger transportation. This dissertation confines itself to consumer goods.

(a) How are the consolidation (sub-) processes at the BBP configured?
(b) How are last mile delivery tours planned and executed?
(c) How is the hand-over process at the point of delivery performed?

This decomposition is meaningful if different perspectives are assumed. The consolidation sub-problem at the beginning of the last mile is a problem of availability and fulfillment strategies (Otto 2008). It is thus a problem of logistics that also benefits from the chain-wide SCM view: should the destinations on the last mile receive supplies from local or rather from central inventories? The sub-problem of transportation is the problem of how shipments are moved on the last mile. Therefore, transportation can be investigated using the functional definition of logistics. The sub-problem of hand-over is a problem of channel design, the distinct way how products are offered to customers. The selection of the points of delivery and the design of the hand-over process there are questions of product characteristics, service features (Schögel 2012, pp. 363–364), and additional services. The distinct services around a product are an “envelop” that consumers value as more or less convenient, reliable, and helpful (Fuller et al. 1993). Therefore, the design of the hand-over process is a marketing problem.

Different last mile problems can be classified by the responsible actor and the type of point of delivery. The last mile can be decomposed into three sub-problems.

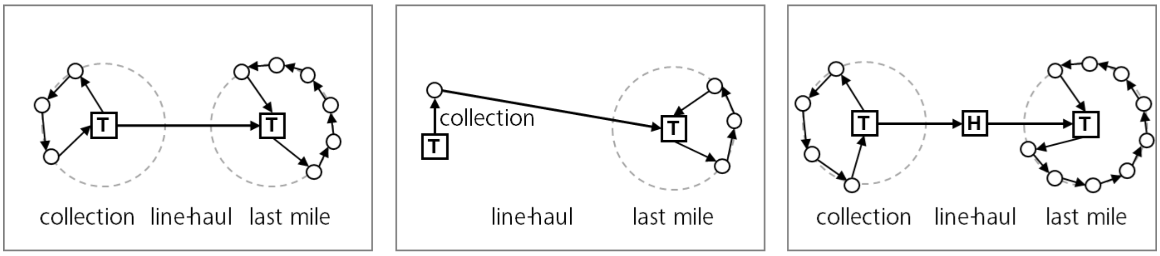
1.3.2 Issues of last mile transportation

It is difficult to operate on the last mile and worthwhile investigating why. Summarizing the issues, the last mile is usually costly, “dirty”, and these two harms should be shared fairly among shippers, stops, and shipments. The following three sub-sections elaborate these issues and their causes.

1.3.2.1 The last mile is expensive

The last mile is the most costly part of distribution because there is less leverage to exploit economies of transportation. The key problems are that delivery tours are composed of a significant level of geographical dispersal of different destinations and the relative small stop sizes in terms of tonnage[27] (Koether 2014, p. 204; Larke et al. 2018). For example, Table 1.2 and Fig. 1.8 dissect the freight costs of three different types of shipments into the cost of the collection tour, the line-haul, and the last mile delivery tour.

[27] Basically the same issues drive the cost of the first mile, Macioszek 2018.



(a) (b) (c)

Fig. 1.8 Examples for freight calculations

The first type of shipment in Fig. 1.8 (a) is a typical 350kg LTL shipment that is transported together with similar shipments in open freight forwarding networks. The terminals are connected via direct lines without a hub in-between. For the calculation, it is assumed that the stop-factor in the collection is 3, which is reasonable because the origins are often manufacturers supplying multiple destinations, as opposed to the stop-factor in the delivery of 1.2. Further assumptions are listed in Table 1.2. The second type in Fig. 1.8 (b) represents one shipping company supplying three destinations in a delivery area. The shipments are quite large LTL shipments of five tons each. It is assumed that the truck collects the three consolidated shipments and transports them directly to the receiving terminal in the destination area. Therefore, there is no collection tour calculated and the approach is included in the line-haul. The delivery tour starts from the terminal[28]. The third type in Fig. 1.8 (c) is a 10kg GF shipment and is loosely based on the case of in-night express (see section 3). In this case, line-haul distance is relatively long because the line-hauling vehicles (vans) travel from the origin terminal to the central hub and back, it is a two-way distance. In this case, the shipment size is small and the stop-factor in the collection is very high. As a result, there are many shipments bundled on a tour-level. A realistic example for this case are auto parts that are used in auto repair shops.

Table 1.2 Freight calculation examples dissected into collection, line-haul, and delivery

		(a)		**(b)**		**(c)**	
Assumptions		ø LTL shipment		Large LTL shipment		GF shipment	
Weight	kg	350		5,000		10	
Line-haul distance	km	350		500		600	
CoDi distance	km	35		35		50	
Stop-factor collection	Sdg	3		3		20	
Stop-factor delivery	Sdg	1.2		1.2		1.2	
Line-haul tonnage	kg	10,000		15,000		1,500	
Line-haul pallets	#	34		30		Not palletized	
Freight costs							
Line-Haul	EUR/LKW	404.51		674.63		339.12	
	EUR/kg	0.04		0.03		0.23	

[28] Daganzo (1987) argues that there is no BBP required in this simple one-to-many case (b), because the truck can simply perform multiple stops in the destination area. Nevertheless, I want to dissect the costs of line-haul from the costs of the last mile and thus assume that there is a break-point in-between line-haul and the last mile.

	EUR/shipment	14.16	31%	168.66	73%	2.26	15%
Collection	EUR/shipment	9.75	21%	-	-	1.12	8%
Last mile	EUR/shipment	21.46	**47%**	62.62	**27%**	11.51	**77%**
Freight	EUR/shipment	45.37	100%	231.28	100%	14.89	100%

The most interesting numbers (bold) in Table 1.2 are the cost of last mile delivery as a percentage of the shipment costs. The percentage varies significantly between the cases. For the typical LTL shipment (a), the last mile accounts for 47% of the costs per shipment, which is remarkable as the line-haul distance is ten times the average delivery distance of 35km. The costs of the large LTL shipment (b) are mainly driven by the line-haul: the distance is almost fifteen times longer than the onward-carriage from the receiving terminal. However, the last mile accounts for 27% of the costs. In the case of the GF shipment (c), the last mile accounts for remarkable 77%. Of course, the collection is very cost efficient due to the stop-factor of 20. However, the two legs of the line-haul are each 300km long and together twelve times the onward-carriage of 50km. Summarizing, the smaller the typical shipment size is the costlier is the last mile delivery of that shipment. This result highlights the relevance of the last mile and the cost issue.

The estimations in the examples above assume a homogeneous delivery process without further restrictions or uncertainty. In reality, there are structural problems at small outlets or homes (Müller and Klaus 2009): vehicle sizes, road access, and parking are restricted; diesel engines are threatened by local regulations in environmental sensitive areas; time slots for delivery to small outlets are narrow due to arbitrary store hours; often there is no docking or unloading equipment at a stop available. The duration of a single delivery can vary significantly due to waiting, searching, securing the load, supporting recipients, or checking back with the back-office.

There are also some ongoing macroeconomic trends that concern the cost issue. These trends and their cost driving effects are discussed in the following.

The shipment structure on the last mile is atomized. There is a growing volume of more than 3.35 bn. parcels in the German CEP market volumes (Bundesverbandes Paket und Expresslogistik e. V. 2018). On the one hand, this ongoing trend to ever-growing CEP volumes enables denser delivery tours having a shorter stop-to-stop distance. On the other hand, the size of individual shipments shrinks year by year: Otto et al. (2018, p. 743) show that the shipment structure of a typical FMCG manufacturer is dominated by small shipments. Also in the case of Amazon, the average number of products per order decreased in Germany from 1.76 in 2004 to 1.33 in 2017[29] (Stüber et al. 2018). Since the number of products per order decreases, the average value per order decreases too. As a result, the distribution costs as a percentage of the shipments' value increase, but the willingness to pay diminishes.

Unattendance diminishes productivity in home-delivery. The productivity of delivery tours in terms of delivered shipments per tour decreases due to customer unattendance. CEP providers need processes to deal with unattended deliveries or to deliver to mobile recipients (Umundum 2015).

[29] Amazon Prime customers order even less items per order: 1.32 in 2017.

Pick-up solutions, which overcome the unattended delivery problem, account for a minor share of only 3%, resp. post offices, as pick-up points, for 10% of the CEP volume in 2017 (Bundesverbandes Paket und Expresslogistik e. V. 2018). Not only private households, also the staff of smaller retail outlets is not always present. While the store personnel is serving customers the waiting time accumulates (Müller and Klaus 2009, p. 146).

Innovative last mile strategies and delivery options drive the transportation costs for two reasons: increasing power of final customers to select the LSP and "cherry picking" of shippers.

"Amazonization" creates an exploding variety of additional services ("Amazonization" Stüber et al. 2018, Hotz and Fost 2016). As an LSP, one has to keep up with these new options in order to stay in business and get one's share of the e-commerce cake: previously the shipper/retailer chose the CEP provider, but nowadays recipients experience the power to choose how, when, where, and by whom they want to be supplied (Umundum 2015). This new technical and organizational complexity drives the difficulty of the coordination function.

Retailers "pick the cherries" for themselves. Large online retailers, with Amazon leading the way, start owning their own rolling assets. Those retailers on the road pick the cherries out of the shipment basket – the dense destinations with great stop-factors. Therefore, CEP LSPs will stick with the "junk" shipments that no one wants to transport. The problem of "junk" shipments is that they do not permit economies of transportation because they have a drop-factor close to 1.0 and a great average stop-to-stop distance. Therefore, neither tours nor stops can achieve high levels of productivity on the last mile anymore.

1.3.2.2 There are emissions on the last mile

Another problem of the last mile transportation is the environmental impact. Edwards et al. (2010) explore the carbon footprint on the last mile. They compare the emitted CO2 from e-commerce home-delivery and from conventional shopping trips. They cannot find an overall emission advantage of one or the other last mile strategy, but home-delivery is likely to emit less. In terms of mileage per item, home-delivery is more efficient than an individual customer travelling to an outlet is. The key driver is the number of items per shopping trip or per home-delivery. For example, even though the number in home-delivery is very small (e. g. 1.4 for books), the pick-up-factor for a car-based shopping trip would need to equal 32 items just to make the break-even (Edwards et al. 2010). "On the way"-shopping that is performed during a trip that is undertaken anyway reduces the total green-house gas (GHG) emissions from last mile distribution. The concept of crowd delivery wants to exploit the omnipresent flow of individual travels (e. g. Devari et al. 2017). The basic idea is to let shipments ride on top of the flow of travelling individuals like rafts that ride on a river.

Kellner (2016b) explores the impact of traffic congestion on the GHG emissions in distribution networks. Traffic congestion increases the GHG emission almost always because congestion leads to detours that avoid congestion and thereby increase total mileage. Further, additional delivery tours are required because the average number of stops per trip decreases due to congestion.

1.3.2.3 The last mile has an allocation problem

There is an allocation problem that arises from the costs and the emissions of last mile delivery. This emission allocation problem is important due to legal reasons. Several national and European committees suggest regulation for the allocation of GHG (Kellner 2016a). The allocation of costs is important in the sales of logistics services and for transfer pricing in freight forwarding cooperations (see section 4).

How should that sort of harms – costs and emissions – be allocated onto shipments, stops, or shippers in road freight transportation? Assuming an LSP is capable to accurately calculate the total costs and the total GHG of a tour, what is the fair share to charge on the individuals? Obviously, shippers and addressees want to get a small share. Therefore, the allocation problem is to find an allocation protocol that calculates the individual share of costs and GHG. An allocation protocol is the underlying calculation scheme that determines how much costs and GHG of a tour serving several customers will cost to each of them (Anshelevich et al. 2008). The criteria of that sort of protocol are accuracy, transparency[30], completeness, individual rationality, consistency, none redundancy, and comparability (Zhu et al. 2014; Kellner 2016a). Another difficulty is to trade off these principles with simplicity. Nevertheless, Zhu et al. (2014) and Kellner (2016a) agree that this ideal cannot be reached weighting all principles equally at once.

1.3.3 Drivers of the shipment cost: the ring-model

Costs play a major role in the course of this dissertation. Therefore, the nature of the cost per shipment on the last mile needs further explanation. In sub-section 1.3.2.1, it has been demonstrated that the last mile accounts of a remarkable share of the costs per shipment. The costs of collection and last mile have been estimated using the ring model. The ring-model estimates the costs of shipments in one-to-many tours (Fleischmann 1998, pp. 67–68). Therefore, it permits sensitivity analyses on the last mile's cost drivers. The purpose of this sub-section is to elaborate the mechanics of last mile costs.

The ring model makes the following assumptions:

- About the tour: every tour is a round trip to equidistant, homogeneous stops that are located concentrically around a central facility (terminal). The CoDi vehicle starts at the terminal, approaches one stop after another, and returns to the terminal. The approach distance and the stop-to-stop distance are the same for all stops.
- About the shipments: the shipments on a tour are homogeneous in terms of size/weight and handling requirements (the time to stop is a linear function of the homogeneous weight). There are no individual service features (like time-windows).
- About the CoDi vehicle: both the size (max. payload) and the allowed driving tour duration are constrained.

[30] Transparency goes hand in hand with a certain level of simplicity. Kellner 2016a uses simplicity and pragmatism synonymously.

- About the cost-types: The usage costs per hour and per km are known and linear.

These assumptions permit comparable estimations of costs and their even allocation to shipments. For the estimation of the costs of a single shipment, the CoDi is filled with identical shipments until either the weight capacity or the tour duration constraint is binding.

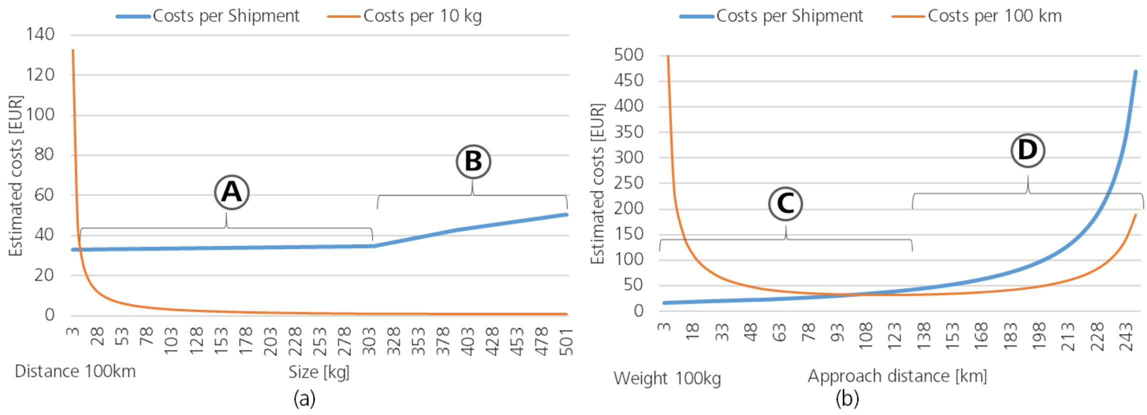


Fig. 1.9 Ring-model estimates of last mile transportation costs using a small CoDi vehicle

In practice, LSPs design their tariffs based on the dimensions (a) size and (b) distance. The graphs in Fig. 1.9 show the estimated costs of tour with a small CoDi truck having a maximal payload of 3.5 tons and a net driving time of 7.75 hours. The costs per shipment are monotonically increasing by size and distance. The graph Fig. 1.9 (a) shows the costs as a function of the shipment size (here the shipment weight is a proxy for size). It is assumed that all stops are 100km away from the terminal (approach distance) and 10km away from one another (stop-to-stop distance). The shipment costs in the segment labeled as **A** have a very small positive slope. In this segment, the CoDi is only partially loaded but the maximal tour duration is fully exhausted. Therefore, additional weight is afflicted with very small marginal cost. Only the stop-duration as a linear function of shipment size increases and thus reduces the overall tour productivity[31]. That kind of tours are "duration dominated" (Hall 1991) and the marginal cost of one additional kg is small. Nevertheless, the orange curve "Costs per kg" is extremely high for small shipments. Since the tour is "duration dominated", the costs of that tour are allocated to very few kg and the small shipments are extremely costly per kg. This is a critical insight, since the majority of all shipments is of rather small size (Otto et al. 2018). The segment labeled as **B** is apparently steeper. At the knee between **A** and **B**, both the duration and the weight capacity are fully exhausted and binding. The slope of segment **B** is steeper because the tour is "weight dominated" (Hall 1991); this means the vehicle's payload is exhausted. Therefore, the marginal cost of one additional kg are high due to the negative impact on tour productivity.

The graph Fig. 1.9 (b) shows the costs as a function of the approach distance from the terminal to the concentric stops. Again, it is assumed that the stop-to-stop distance is 10km. Further, it is assumed the shipments have an equal weight of 100kg. There is an important trade-off that is

[31] Tour productivity is a measure of the output efficiency. Herein, the number of stops per tour is the measure of tour productivity.

visualized by the orange curve "Costs per 100 km": if the non-productive[32] approach distance increases, there are two antagonizing effects.

(1) On the one hand, the total costs per km decrease hyperbolically: there is a cost degression effect. As the non-productive approach distance increases the total costs are allocated to more and more km. In addition, the driving velocity increases and more and more km are covered per hour.
(2) On the other hand, the tour productivity decreases because a larger share of the net driving time is used on the non-productive approach.

The segment labeled as **C** has very short approach distances. As a result, the CoDi vehicle can be fully loaded and the tour productivity is very good. In segment **C**, the costs per 100km decrease due to the cost degression effect: the personnel and vehicle costs per hour are allocated onto more and more kilometers. The degression effect dominates in segment **C**. At the turning point of the orange curve "Costs per 100km" between the segments **C** and **D**, this trade-off tips to the other side. In segment **D**, the decreasing productivity effect dominates.

From Fig. 1.9, the following insights are drawn:

- The costliest shipments per kg are the smallest shipments. The most efficient shipments in terms of size are large shipments: the larger the better.
- The marginal cost per kg are small on "duration dominated" tours and vice versa.
- For different approach distances, there is a trade-off between the cost degression effect and the decreased productivity effect. Therefore, the most efficient shipments in terms of distance are (depending on the vehicle and cost coefficients) in the middle range.

The stop-to-stop distance is another major driver of tour productivity. If stops are close to one another (the tour density is higher), then there is more time to make more stops. Fig. 1.10 shows the estimated costs per shipment as a function of the stop-to-stop distance. The underlying calculations assume an approach distance of 100km and shipment size of 100kg. The orange dashed curve represents the costs for a constant velocity between stops of $vs = 33 \, {}^{km}\!/\!_{h}$. The blue curve represents the costs for a linear ascending vs . If the stop-to-stop distance increases, the CoDi truck can accelerate and thus the increase of the driving time between stops is disproportionately smaller. The difference between the orange and the blue curve represents the effect of this acceleration.

[32] „Non-productive approach" means that there is "dead-mileage" (van Oudheusden et al. 1999, p. 524) traveled without servicing stops. As a result, the tour productivity degrades.

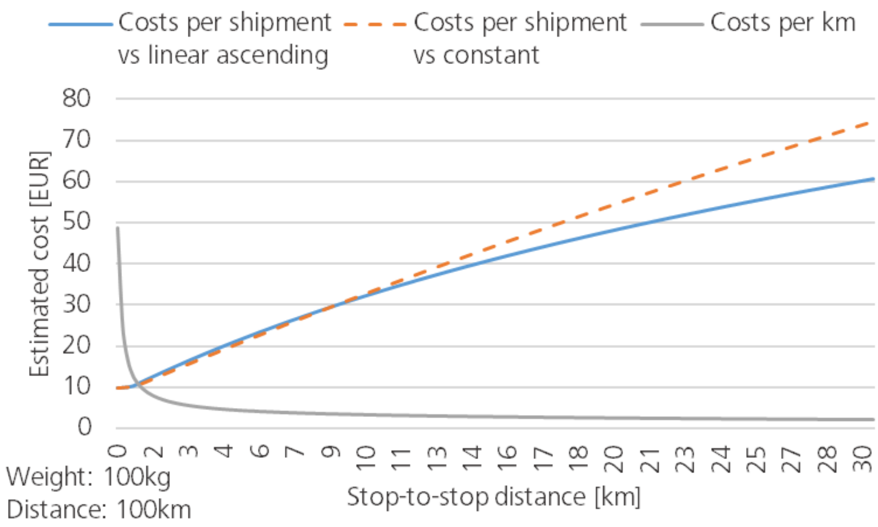


Fig. 1.10 Costs per shipment as a function of the stop-to-stop distance

The hyperbolic grey curve represents the costs per km from stop to stop. The marginal cost per km are monotonically decreasing. By implication, there is great potential for savings from shortening the stop-to-stop distance (improve the tour density) because the marginal savings are increasing. For example, if the stop-to-stop distance of 10km can be shortened by 1km, then EUR 1.97 per shipment can be saved. For comparison, in order to achieve the equivalent savings, the approach distance would need to be shortened by 7.6km. If the tour can be densified by yet another 1km, then EUR 2.06 can be saved per shipment.

The costs per delivered shipment depend strongly on the average number of shipments per stop – the drop factor. A densification of stops (higher drop factor) reduces the costs per shipment significantly. Fig. 1.11 shows this relation using the same example as before. The costs per stop are slightly increasing since the delivered tonnage is proportional to the drop factor and thus the variable stop time increases. However, the degression effect overcompensates this such that densification has great potential for savings.

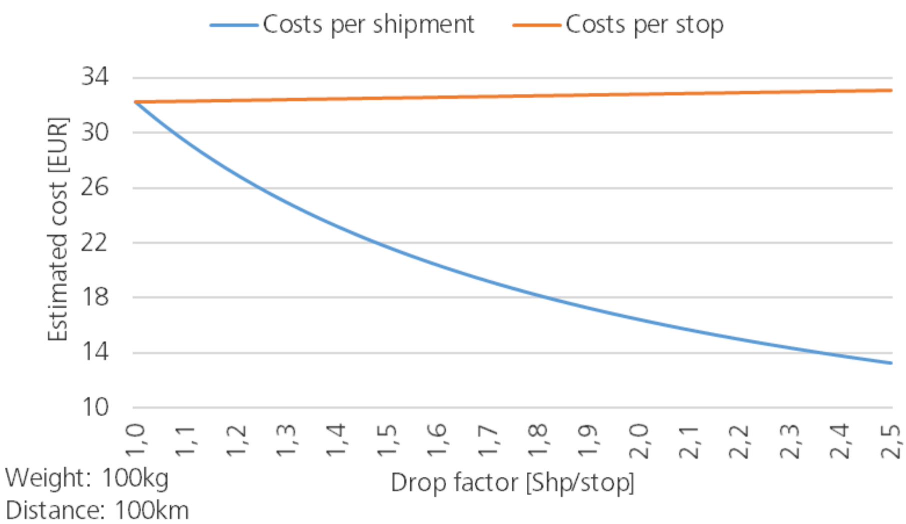


Fig. 1.11 Costs per shipment as a function of the drop factor

Summarizing, the costs per shipment on the last mile depend on (1) the shipments' size and distance and (2) on the LSP's capability to bundle shipments together in a way that densifies tours and stops and thus improves the tour productivity through greater drop factors (stop densification) and shorter stop-to-stop distances (tour densification).

The shipment structure is subject to the replenishing and ordering behaviour of manufacturers and retailers. The acquisition of suitable shipments that densify tours is a marketing task (sub-section 1.1.1). The chain wide coordination of ordering and the establishment of cooperative planning with the goal of higher densification on the last mile are tasks of the SCM perspective (sub-section 1.1.3). The bundling of existing shipments can be managed and is a logistics planning and optimization problem (sub-section 1.1.2).

1.3.4 Decision Problems of last mile distribution-logistics

Having an understanding of what the last mile is and what the major issues of the last mile problem are, this sub-section outlines decision problems of last mile distribution from a logistics perspective. Decision problems of channel design and inter-channel coordination (marketing) and the inter-organizational, boundary-crossing relation of production and consumption (SCM) are important fields of research. Nevertheless, they are not discussed herein. Fig. 1.12 gives an overview of typical decision problems which occur on the last mile of distribution systems. The decision problems are categorized by the decision horizon (strategic, tactical, and operative) and by the three sub-problems of the last mile problem (consolidation, transportation, hand-over, see Fig. 1.7). However, some decision problems that touch on multiple sub-problems are insinuated horizontally across borders.

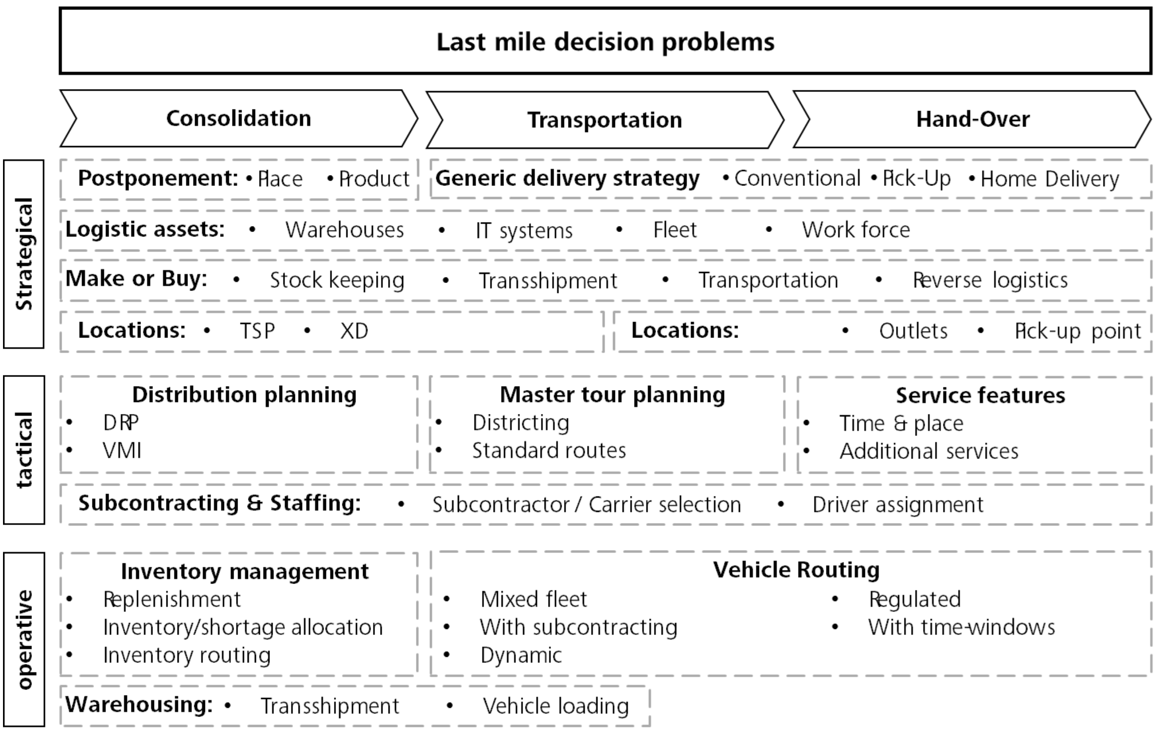


Fig. 1.12 Last mile distribution-logistics decision problems

1.3.4.1 Strategical decisions

The **generic delivery strategy** set a framework for many hierarchically subsequent decisions. The strategic alternatives of last mile delivery concern the question, where the point of delivery is located in the spectrum between large outlets and home delivery (see sub-section 1.3.1.2 Systematization). This decision makes a trade-off between the service quality and the customer involvement in order to exploit revenue and cost drivers (Chopra 2003). The last mile gets more difficult and thus costly, the closer the point of delivery moves to the customer's home because the customer gets less and less involved, taking ownership of less activities. On the upside, individualized service features (e. g. delivery time-window) and the capability to yield additional services (e. g. setup) in home delivery permit exploitation of additional revenue drivers (Chopra 2003). Fuller et al. (1993) argue that products are always the offered things and their enveloping services. The services on the last mile are presumed to create value for customers through convenience, reliability, and support. However, there is no "one-size-fits-all". Such an approach would likely under-serve sophisticated customers with special requirements and over-serve (and over-charge) basic customers with general needs. The fit of different delivery strategies depends (a) on the product characteristics of FMCG and SMCG, (b) on the additional services (Schögel 2012, pp. 363–364) and (c) on delivery service features.

(a) The characteristics of FMCG and SMCG differ. Relevant characteristics are the frequency of purchase, the value density, the average margin, the substitutability by competitive brands, and the customers' willingness to switch brands. The characteristics of FMCG requires immediate availability, whereas SMCG tolerate some lead time. The product characteristics determine a product's suitability to postpone the last mile delivery until the customer really makes an order (Pagh and Cooper 1998). Consequently, the conventional distribution of FMCG ends at the "brick and mortar" outlet. Vice versa, the last mile distribution of SMCG is postponed and they are delivered to pick-up points (click&collect) or home.
(b) The opportunities to perform additional services differ between the strategies[33]. In conventional distribution, services like shelf stowing are performed for the retailer. Additional services are an important reason for manufacturers to use a DSD strategy (see sub-section 1.2.2). In home delivery, services are performed for the final customer. Since additional services always come with additional costs, these services are only justifiable if there is additional value for the customer that is appreciated and can be developed as a competitive advantage.
(c) Service features on the last mile are duration of delivery, time-window of delivery, spread (size) of the delivery area, and reverse logistics (Kuhn et al. 2018). These service features require different configurations of the routing and the hand-over processes. The more sophisticated the service features are, the costlier the strategy is. Highly individualized service features are justifiable for high-value and high-margin products.

[33] I discuss additional services that are performed during the delivery by the same personnel, like setup of white goods or disposal of old devices. Other services that are performed during the demand acquisition, such as advising by sales people, e. g. Stihl handheld power equipment, or financing services are important considerations. Demand acquisition services are a marketing problem. A strategy that combines that kind of demand acquisition services with home delivery fulfillment is the "showroom", see Bell et al. 2018. Those integrated channels are not discussed herein.

These three decision variables are explained in

Table 1.3 with respect to FMCG and SMCG.

Table 1.3 Typical decision variables of last mile strategies

	Conventional distribution	**Home Delivery**
Product Characteristics	The demand for FMCG is relatively stable and their value density is low. The substitutability is high and customers' brand loyalty is often low. Therefore, FMCG are distributed into local retail outlets speculatively.	The demand for SMCG is often unstable and their value density is high. The complexity of the financial risk of purchasing an SMCG is high, therefore customers' willingness to switch brand is low. As a result, SMCG are stored centrally and distributed to customers' orders.
Additional Services	FMCG handling is easy. There is minor need for additional services. The distributor performs services for the retailer.	SMCG have characteristics (size, complexity) that require additional services. The distributor performs additional services for the final customer.
Service features	Distribution of FMCG should be cost-efficient, thus service features are highly standardized.	SMCG provide high margins and the customer has a high willingness to pay. High service quality and individualization build customer satisfaction.

The **postponement** decision is about speculative versus order-driven distribution and closely linked to the generic distribution strategy. This decision is a classical SCM topic (Pagh and Cooper 1998). On the last mile, there is a minor postponement problem. Postponement of placing inventories means if products are already on the last mile, when they are demanded or not. For example, delivery of products within very few hours requires urban warehousing. If that sort of extremely fast delivery service is offered, a highly speculative distribution of products to many decentralized urban warehouses is mandatory. Product postponement means the final packaging, configuration, or shaping of an otherwise finished good (cf. Brown et al. 2000). Cutting of wooden planks, packaging of fresh meat products, arrangement of flowers, or final configuration of consumer electronics are examples of product postponement on the last mile. Postponement is a set of Boolean yes or no decisions: "Do we offer ready packed meat or do we cut and pack it up?", "Do we configure a device on demand, or do we simply offer the already finalized device?"

Make or Buy (MoB) is a major decision of strategic management and has a special relevance for last mile distribution. The alternatives are to make a process in-house or to delegate it to an LSP. Typically, the processes in question are transportation, transshipment, stock keeping, and reverse logistics. Many manufacturers and retailers decide to outsource the distribution completely to an intermediary distributor (Chopra 2003). The make or buy decision requires a careful analysis of costs and revenue drivers but also of other criteria (Bretzke 1989; Cooper et al. 1997; Chopra 2003). Most often outsourcing is beneficial due to lower expenditures: the costs per piece are lower due to better fixed cost degression of an LSP, opportunity costs are lower due to less capital employed, and most often there is arbitrage from lower LSP wages to exploit (Bretzke 1989). On the opposite, if a service is "bought", the transfer, the coordination, and the control of the service is expansive. The concept of transaction costs tries to capture the additional costs from outsourcing (Picot et al. 2010). Picot et al. (2010) argue that the application of advanced IT systems decreases the transaction costs and makes outsourcing more attractive.

Closely related to the MoB decision are investment decision problems with respect to **logistics assets**. The decision to make last mile distribution in-house requires a decision to invest. Typical investment assets are the already discussed locations, the fleet, the human work force, and the IT systems. Investment decisions are typically modelled as net present values. The present value method is a method that supports the long-term selection of investment alternatives. A related concept is the economic value added, which is an indicator of the monetary value created (Lambert and Burduroglu 2000). Decision problems are the heterogeneous fleet sizing, the investment decision in logistics real estate, the sizing of the work force, hire and fire problems, and IT investments.

Location decisions determine the physical nodes of last mile distribution. On the last mile, there are transshipment points or XD which serve as the BBP to approach many small destinations. Depending on the last mile strategy, the system deploys outlets and pick-up points, where consumers get involved in the distribution process. Such locations require investments in real estate. The network of pick-up points must cover a demand area in a suitable way. The pick-up location problem is a typical network design problem and trades off warehousing and transportation costs and service levels (cf. Kellner et al. 2013)

1.3.4.2 Tactical decisions

Master tour planning is the clustering of potential destinations in a demand area such that every transportation requisition is automatically assigned to one unique master tour. Districting is the problem of partitioning a larger geographic service area into smaller districts in such a way that the resulting districting plan is acceptable according to some planning criteria (Shirabe 2009; Kalcsics 2015; Brabänder 2018). A district is a contiguous part of the service area that is always served by the same vehicle/s and/or employee/s. A standard or fixed tour is an exact schedule of frequently supplied destinations. A standard tour is assigned to one vehicle that frequently operates the identical schedule. The major difference between districts and standard tours is the stochasticity of the number and exact location of destinations. The stochastic destinations that occur in a district are routed in daily operations. Standard tours have a schedule of mostly fixed destinations with relatively stable demand (Schmidthöfer 2004). Standard tours are often used in the distribution of letter mail and newspapers. Districting is often used in groupage freight and CEP distribution.

Distribution planning is the determination of volumes in a hierarchical structure. Distribution requirements planning (DRP) is a planning concept to match finished goods stock and independent demand in a multi-echelon distribution system. DRP is part of the demand planning and integrates demand forecasting with inventory management and master production scheduling. This integration takes place in the periodic stock levels of RWHs, CWHs, and the production lots. Inventory requirements are derived from forecasted demand, distribution requirements are derived from inventory requirements (Vollmann et al. 2005). For the last mile, DRP gathers the final customers' data and aggregates it on the closest institutional stock keeping location. Therefore, DRP determines the planned stock levels of products that are immediately available for last mile distribution. This comprises safety stocks. The question is, whether there should be safety stocks close to the customer or not. Centralization of safety stocks reduces inventory cost but decentralization improves availability (Tempelmeier 2015). Most contributions on the placement of safety stocks recommend an "all or nothing strategy", which comes down to the decision if there should be safety stock on the last mile or not (Simpson 1958; Inderfurth and Minner 1998). VMI is a replenishment concept which means that a manufacturer makes the DRP for the last stock keeping facility and thus for the last mile (Tyan and Wee 2003). VMI is deployed to mitigate the bullwhip effect in supply chains (Lee et al. 1997). The last mile is the first stage of implementation for this concept. From the logistics perspective, VMI needs an interface between the vendor and the distributor. The distributor on the last mile needs to coordinate this interface.

The **service features** that are offered are a tactical decision with the goal of customer satisfaction. Service features can be changed relatively fast (mid-term horizon) in order to readapt to environmental changes quickly. This decision makes a trade-off between costly individualization of features and exploitation of higher willingness to pay for superior service quality.

Subcontracting & staffing is the selection of subcontractors and the assignment of tasks to either the own work force or subcontractors. The selection of a subcontractor is a multi-criteria decision problem. Therefore, the LSP selection has been tackled with the analytic hierarchy/network process (Jharkharia and Shankar 2007). Menon et al. (1998) found that LSP selection is driven by the perceived performance and capability of the LSP. The rates of the LSP are of inferior importance: "Performance and quality requirements must be met before serious rate discussions can occur" (Menon et al. 1998, p. 130). The selection process is similar to the outsourcing of other industrial services: identify own requirements, outline tasks to outsource, identify potential providers, evaluate providers, select the most suited provider, and negotiate the service agreement (Menon et al. 1998). Staffing of own and subcontracted resources comes down to a resource allocation problem: all tasks have to be done and resources should be allocated onto tasks in an efficient way.

1.3.4.3 Operative decisions

Inventory management on the last mile is the operational ordering of replenishment lots for the warehouse closest to the customer. The DRP has determined the volumes that are planned for a demand area. In the short-term, actual customer orders arrive and the plan is adapted to the actuals. The inventory routing problem is the repeated distribution of goods to many customers, where the distributor is responsible for the customers' replenishment. For example, a vendor in a VMI agreement with a chain of discounter outlets has to make inventory routing decisions. Therefore, inventory routing integrates the inventory lot sizing and the vehicle routing decision (Campbell et al. 1998). On the last mile, shortages, which are then filled from the central warehouse, may occur. In that case, the service promise cannot be met anymore. Therefore, the shortfall is allocated to customers: which customer should receive a scarce product and which should wait. Fischer (2001) proposes a linear model which minimizes to cost of allocation of inventories to actual demands. He measures the cost of allocation with four coefficients: temporal, priority, profitability, and historic revenue of an actual order.

Vehicle routing on the last mile is daily business. The routing of a single vehicle through many points and return to the depot is known as the travelling salesman problem (Applegate et al. 2011). The assignment of points to different vehicles and routing them is known as the vehicle routing problem (Golden and Assad 1988). There are many extensions to this problem. Last mile routing faces problems like a heterogeneous (mixed) fleet (Yaman 2006), time-windows (Schneider 2016), driving time regulations (Meyer and Kopfer 2008), city regulations (Cattaruzza et al. 2017), the coordination of the own fleet, common carriers (Krajewska and Kopfer 2009), and subcontracted fleet (Gahm et al. 2017).

Warehouse operations require decision-making on the transshipment of goods inside a transshipment point. Milk run operations transfer goods from inbound vehicles to the re-assembling and packaging station and further to outbound vehicles. Homogeneous palettes from FTL shipments are opened and re-assorted to mixed palettes for outlets. In CEP services, inbound shipments are unbundled and sorted for the distribution districts or pick-up points. The assembled shipments are palletized, packed or wrapped, and labeled. On the outbound end of warehousing, shipments are bundled to loads and loaded into CoDi vehicles. Vehicles are loaded according to the vehicle routes and product requirements, in a way that enables efficient unloading and delivering.

1.4 Outline

This dissertation focuses on model-based decision support of last mile transportation. The following sub-section defines what a model and what model-based decision support are. The remainder of this section outlines the three focus papers of chapter 2-4 of this dissertation.

1.4.1 Scope of this dissertation

The scope of this dissertation is limited. The addressees are LSPs, hence the last mile problems are investigated from the perspective of an LSP. Typical LSPs that are addressed in this dissertation are GF and CEP service providers. The focus papers of this dissertation investigate last mile problems with many small shipments, which are either small B2B or B2C shipments. Last mile problems of manufacturers and retailers and large B2B shipments are out of scope (see the systematization in sub-section 1.3.1.2). The object of the investigation is the last mile transportation of consumer goods. Therefore, the consolidation processes at the BBP and the hand-over processes at the points of delivery are out of scope (see Fig. 1.7). The cost issue of the last mile is the focus of this dissertation. The other issues are not fully out of scope but side issues. The three distinct focuses are outlined in the course of this sub-section. The applied methods in these papers are model-based approaches (next sub-section) to analyze decision problems of last mile transportation (see Fig. 1.12). Other approaches to deal with the last mile issues are not investigated herein. The outlook on further research in sub-section 5.2.4 proposes a non-exhaustive list of other approaches.

1.4.2 Model-based decision support

This dissertation takes a quantitative analytical – more precisely a model-based – approach to decision support. Decision support systems (DSS) can be classified as either model-based or data-driven. In practice, the key difference is the implementation of model-based approaches in business analytics systems and of data-driven approaches in business intelligence systems (Fink et al. 2014). Business intelligence means an information system that collects, holds, connects, and analyses business-related data. Business analytics supports decision-making by modelling decision-making processes and determining decisions using data (from several sources including business intelligence) and application-specific methodology (Hansen et al. 2015). Due to this relation, business analytics requires business intelligence as a source of high quality data for the purpose of decision-making.

What is a model? A model has a paradigm[34] and a purpose. There is some real-world original to be modelled and the model makes inherent judgements about the relevance of aspects of the original. The model's purpose is used to judge which aspects to flag as relevant and which not. Then, having made this judgment, a model makes a projection of the real-world original into the paradigm's domain, dropping the non-relevant aspects. A model is therefore a reduced version of the original. The paradigm is the character of the projection and implies that there is a guiding reference establishing an association: "it is like this". This reference establishes the form of the selected aspects in the model: a model is something formal. This means that the model-building process is an act of guided formalization. The outcome (the model) is always formal, rule bound, more explicit, and more exact than the real world which has informal, arbitrary, and less relevant aspects (Morgan 2012, pp. 19–20). The domain of the model's paradigm adds "abundant" aspects to the relevant aspects (Stachowiak 1973). Those abundant aspects augment the otherwise reduced projection such that a model is more than a formalized stripped original. Abundant aspects are what can be observed in addition to what has been labeled as relevant aspects. For example, the concept of the BBP is an abundant aspect of distribution-logistics as a system in which CG flow: multiple sequential transportation and sorting processes. The BBP divides that sequence into the consolidated "few-to-few" and the last mile "one-to-many" part.

Having said all this, a model is defined as a purposeful[35] and reduced projection (Raffée 1974, p. 57) of an original real world subject into the domain of its paradigm (Stachowiak 1973; Otto 2002; Hansen et al. 2015).

In the model-building process, there are always multiple imaginable paradigms of the same original. A model-builder views this original "like something metaphorical else". For example, the model-builder modelling the last mile may view it like:

a. A graph where goods flow between vertices along edges.
b. A geographical map that represents a service area.
c. A hydrodynamic system in which flows are quantified by physical characteristics.[36]

---

[34] A paradigm is a conceptional blueprint or pattern of thinking.
[35] Stachowiak 1973 uses the term "pragmatic" instead of purposeful.
[36] Klaus and Sheffi 2012 propose the hydrodynamic flow analogy for the third definition of logistics.

d. A set of channels that make goods and services available to markets.
e. A process map of processes connected by information and material inputs and outputs.

As a result, the chosen paradigm of a model predetermines what can be learned from that model. Often one paradigm makes implicit assumptions that another paradigm recognizes as decision variables and vice versa. All models omit some aspects of reality, which are not perceived relevant for the model's purpose. For this reason, the quote by Box et al. (2005) "All models are wrong; some models are useful", is famous. Daskin (2010) further justifies the wrongness of models: they provide insights in different aspects of decision problems but may not regard other aspects. Furthermore, the abundant aspects diverge the model and its original even more – but in a purposeful way. The user of a model can gain new insights about the original through observing the abundant aspects of the model. Therefore, the purpose and scope of a model need to be clear. Wrongness simply means that models are not comprehensively true, it does not mean models are all wrong. Decision-makers can apply or formulate models that support them for a specific purpose that may require only a reduced snippet of the real world. Overall, models support decision-making but do not replace the decision-maker. Developing this thought about the wrongness and helpfulness of models even further, Pratchett et al. (2013) described the "lie-to-children" understanding of models. This educational concept highlights both the simplicity of models as well as the endeavor to reach more sophistication: "Science's models are not *true*, and that's exactly what makes them useful. They tell simple stories that our minds can grasp. They are lies-to-children, simplified teaching stories, and none the worse for that. The progress of science consists of telling ever more convincing lies to ever more sophisticated children."

1.4.3 Outline and placement of the three journal papers

Last mile transportation is a major research topic in the field of distribution. Due to the outlined issues and decision problems on the last mile, decision-makers demand decision support models. Fink et al. (2014) identify three challenges to decision support models for real-world use. First, the integration of decisions in manufacturing and service networks should be tackled. Second, robustness is identified as a favorable property of decision support models in order to deal with uncertainty. Third, the acceleration of the model-building process itself is a challenge.

In this dissertation, three major problems of decision-making in last mile transportation are selected, investigated, and published as independent papers. The following Table 1.4 summarizes the three papers. The content and research contributions are then outlined in the subsequent subsections.

Table 1.4 Bibliographic summary of the three focus papers

	Paper 1	**Paper 2**	**Paper 3**
Chapter	Chapter 2	Chapter 3	Chapter 4
Title	Vehicle routing with private fleet, multiple common carriers offering volume discounts, and rental options	Distribution Districting – the case of in-night express services	Bringing infrastructure into pricing in road freight transportation – A measuring concept based on navigation service data
Authors (estimated contribution)	Christian Gahm (45%) Christian Brabänder (35%) Axel Tuma (20%)	Christian Brabänder (100%)	Florian Kellner (45%) Andreas Otto (30%) Christian Brabänder (25%)

Research questions	*Research Question 1*: How can the VRPPCdR be formulated as a linear model? *Research Question 2*: Are solution heuristics based on Variable Neighborhood Search suitable for the VRPPCdR? *Research Question 3*: Do many rental and outsourcing options grant superior cost reductions?	*Research Question 4*: How to model the distribution districting problem using an IP? *Research Question 5*: Do solutions for the DDP from an MIP model perform well in practice?	*Research Question 6*: How can infrastructural differences get measured? *Research Question 7*: How can infrastructural differences be translated into cost differences?
Methodology	Mixed Integer Program and Variable Neighborhood Search	Case study applying an Integer Programming formulation	"Big data" approach for measuring cost differences
Data	Test data sets from literature and modified scenarios	Shipment data from Kiessling-Spedition, geographical data from OSM, navigation data from Google Maps	Shipment data from „FA" and navigation data from TomTom
Addressee	Logistics service providers who distribute consumer goods on the last mile		
Status	Published		
Bibliographic data	Transportation Research Part E: Logistics and Transportation Review, 2017, Vol. 97: pp. 192-216.	Journal of Service Management Research, 2018, Vol. 2, Issue 4: pp. 36-52.	Transportation Research Procedia, 2017, Vol. 25: pp. 794-805.

The three focus papers contribute to decision-making in last mile distribution-logistics. Fig. 1.13 positions the three papers in the context of the decision problems that are elaborated in sub-section 1.3.4. First, the VRPPCdR contributes to the operative vehicle routing problem and considers a mixed fleet, even rental options, and the subcontracting options. The second paper discusses districting in a last mile context and is positioned in the tactical master tour planning. The third paper brings infrastructure into pricing and hence contributes to decision support models across the planning horizons. The cost differences per km can be used as inputs of decision support models for location, master tour planning, and vehicle routing. Eventually, this third paper does not propose a decision support model, but a concept to collect and estimate better input data. Better data makes models a bit more sophisticated, less wrong, more robust.

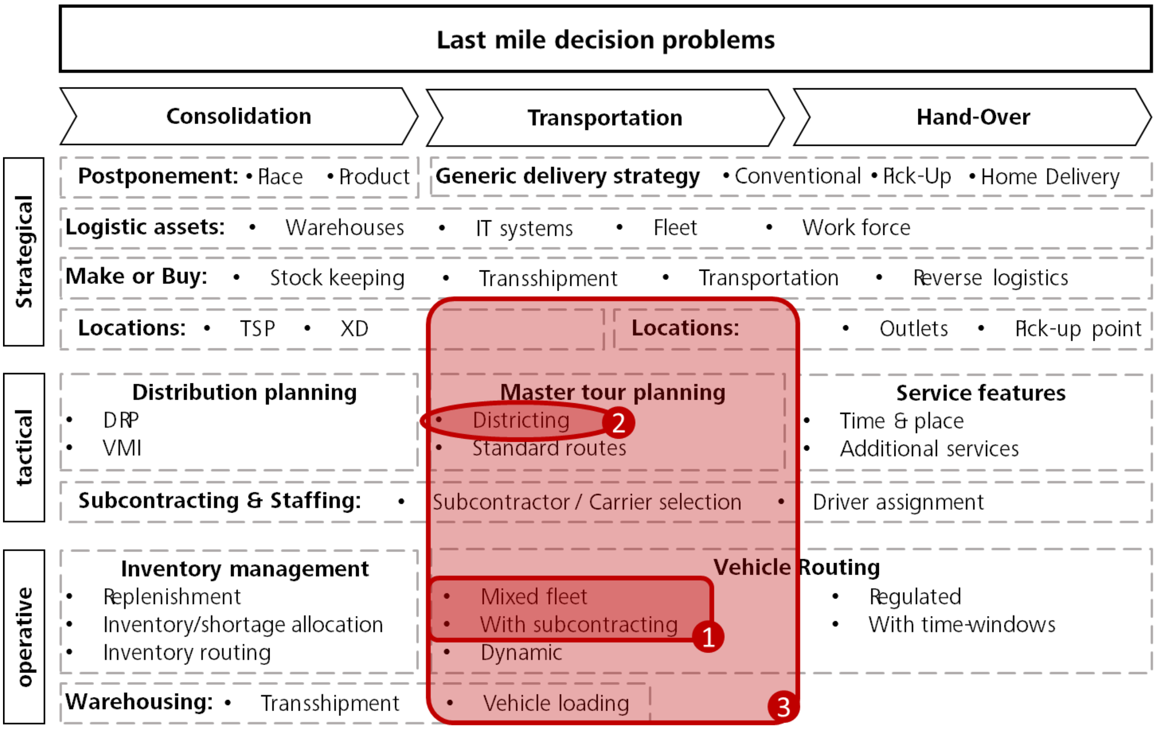


Fig. 1.13 Placement of the three papers within the last mile decision problems

1.4.4 Focus 1: a decision support model for vehicle routing and carrier selection

The first focus investigates the vehicle routing problem with private fleet and common carriers (VRPPC) and extends it. The last mile problem here is of general nature. There is a one-to-many transportation problem where the BBP is a depot. From the depot many shipments are transported to many destinations in the area. The decision-maker has different fulfilment options for the transportation: (a) the mixed private fleet, (b) rental, and (c) subcontracting options. The VRPPC combines the VRP and the assignment of shipments to subcontractors. The assignment decision is essentially the "cherry picking" decision of selecting those shipments that enable great tour-densification and savings and assign them to the private fleet. Every shipment is assigned to either a vehicle of the own fleet, or a rented vehicle, or a common carrier. There are two rental options considered: a time-based and a mileage-based charge. The shipments that are assigned to own or rented vehicles are scheduled. Outsourcing to an LSP assumes that the freight rates have been negotiated in long-term contracts. The freight rate is modelled as a concave function of the size of the shipments. Potential volume discounts are included. The extended problem is called the vehicle routing problem with private fleet, multiple common carriers offering volume discounts, and rental options (VRPPCdR). It combines three interdependent operational decisions: selecting common carriers for subcontracting, clustering shipments in tours and in subcontracted bundles, scheduling the routes of own and rented vehicles. In this paper, this combined decision problem is modelled.

This is an operative optimization problem of contract LSPs who distribute a great number of shipments in the area. The analysis of a typical shipment structure shows that the majority of shipments is small and dispersed. The major share of total weight and volume is in a minor share of the shipments (Otto et al. 2018). The large number of small shipments does not enable great economies of transportation. Therefore, small shipments that do not contribute to the densification of tours are subcontracted.

With respect to the research agenda by Fink et al. (2014), this paper contributes to the integration of service networks through the combination of independent decisions. The first research question (RQ) of the first paper is:

RQ 1: How can the VRPPCdR be formulated as a linear model?

As the VRP and VRPPC are modelled as (linear) mixed integer programs (MIP), also the VRPPCdR is modelled as an MIP. The objective is the minimization of the total cost of the own fleet, rental charges, and outsourcing charges. The constraints of the model are adopted from literature or newly introduced. The outsourcing of shipments with common carriers turns out to be non-linear. Therefore, a linearization is developed and applied. The subtour elimination constraints of Miller et al. (1960) and Yaman (2006) are adapted to deal with the new extensions.

The second RQ is:

RQ 2: Are solution heuristics based on Variable Neighborhood Search suitable for the VRPPCdR?

In order to solve the linearized model, the MIP is at first solved with the CPLEX solver. Literature shows that Variable Neighborhood Search (VNS) is capable to solve large routing problems (Kytöjoki et al. 2007). Therefore, three heuristics based on VNS are proposed, implemented, and tested. The variant called randomized Variable Neighborhood Descent (RVND) delivers the best mean solution quality for the test instances. Comparing the VNS heuristics with the CPLEX solution after 50 hours, the heuristics deliver improved solution quality in extraordinarily small computation time for larger problem instances.

In order to draw economical insights from the model, several scenarios with various combinations of rental and outsourcing options are computed in order to answer the third RQ.

RQ 3: Do many rental and outsourcing options grant superior cost reductions?

The contributions of this paper are the combination of routing and subcontracting as an extension to the VRPPC and the adaptation of linear formulations to deal with them. Second, the three new VNS variants are compared using the VRPPCdR as an example. Third, a scenario analysis explores the economic benefits of having multiple rental and outsourcing options, instead of one common carrier.

1.4.5 Focus 2: a decision support model for the distribution districting

This paper investigates the tactical distribution districting problem (DDP). Districting in the last mile distribution context is a special case of districting problems (Kalcsics 2015). Districting exploits the risk pooling principle through the aggregation of elementary demand units (herein, zip codes) on a district level. In distribution, there is some terminal in the service area that serves as the BBP. There, line-hauled, consolidated shipments are re-consolidated onto delivery tours. Every tour bundles many destinations within one distribution district. The districting problem is how to partition the service area into distribution districts that enable good delivery tours through many destinations inside the district borders.

CEP services that deliver into a larger area from a regional terminal are natural applications of the DDP. In a broader scope, distribution districting also addresses network planners at a larger geographical scale:

- Partitioning of an FMCG manufacturer's retail customer into districts for DSD
- Partitioning of a GF alliance's national territory into its members' service areas
- Partitioning of a CEP network's service area into terminal areas

The major difference between districting a national network into service areas and districting a service area into tour areas is the selection of appropriate activity measures.

The shape of districts is fixed over a tactical planning horizon in order to gain benefits of consistency: driver and customer satisfaction, driver familiarity and learning, reduced error-rates and improved process times, and reduced overhead costs. The exploitation of consistency requires a certain level of robustness. From the perspective of system theory (Ulrich 1968), robustness is the capability of a system to perform its processes and function accurately in the face of unexpected and unfavorable environmental incidents. In short, robustness is the insensitivity of a system's performance to externalities (Schneeweiß and Kühn 1990). Districting makes the distribution system more robust. This paper highlights this achievement and hence contributes to the development of decision support models that incorporate aspects of robustness (Fink et al. 2014).

Shirabe (2009) proposes exact mathematical formulations for the general districting problem. He adopts the net flow paradigm to consider the contiguity of models explicitly. The first RQ of the second paper is thus:

RQ 4: **How to model the distribution districting problem using an IP?**

The paper combines Shirabe's net flow approach with an advanced estimator of tour duration from Schmidthöfer (2004). In addition, the flow decision variables are used to model a new proxy for the compactness of districts which is used in the model's objective function.

In order to evaluate the practical applicability, the model is solved with the CPLEX solver and applied to a real-world distribution case. The second research question of this paper is the following.

RQ 5: Do solutions for the DDP from an MIP model perform well in practice?

An online routing tool is set up to perform daily vehicle routing through the destinations in the calculated districts. The results are evaluated to be of good quality. The IP solution is then compared with two alternative solutions. An adapted savings algorithm (Clarke and Wright 1964) and a rule-based practitioner's solution approach are presented. While the adapted savings algorithm does not produce compatible solutions, the practitioner's approach performs almost as good as the IP approach.

The theoretical contributions of the second focus of this dissertations are threefold. First, an IP formulation for the DDP is presented. Second, the model is combined with an advanced estimator of travel time. Third, the in-night distribution case study demonstrates the applicability. One major implication is the superior robustness of the IP solution in the face of spiking volumes. This is a result of the optimized compact shape of the districts.

1.4.6 Focus 3: a data driven approach that brings infrastructure into freight pricing

LSPs who operate in different districts, cities, or regions have different productivities. There are different drivers of these differences. In the ring-model (sub-section 1.3.3), differences originate from the shipment structure: shipment size, drop-factor, stop-to-stop distance, and the approach distance are the key indicators of those differences in the shipment structure (Fleischmann 1998). Therefore, an LSP's productivity depends on the shipment structure and on the LSP's assets and capabilities.

This paper focuses on a third driver of an LSP's productivity: infrastructural differences affect the average driving velocity in an area. Frequent congestion, natural obstacles, narrow streets in old city quarters, one-way streets, and driving speed limitations on the last mile are examples for causes of low productivity. That sort of infrastructural barriers cause reduced driving speed and long detours. Everything else being equal, the stop-to-stop duration of two LSPs in different regions can deviate. Different subcontractors in different districts may have different productivities, different partners in a GF alliance may have different fulfillment cost, different outlets in different regions may be more or less costly to replenish. Such differences are out of the hands of an LSP, but they affect the productivity regardless. The third research paper investigates this problem. The first research question is:

RQ 6: How can infrastructural differences get measured?

A big data approach based on online navigational data is proposed. A big amount of driving times and distances is collected online from the service area of five LSPs of a German GF cooperation. It is important to measure driving times and distances where the destinations really are, instead of using general city indices.

Infrastructural differences affect the allocation of costs in logistics cooperations. Especially the calculation of fair transfer prices is important for the stability of the cooperation. Therefore, the second RQ in this paper is:

RQ 7: How can infrastructural differences be translated into cost differences?

The goal is to estimate the average cost per km on the shortest way using navigational data. This key indicator isolates the infrastructural differences from differences within the shipment structure and differently productive assets. The online sourcing of data ensures a transparent calculation. Through the measurement and incorporation of the infrastructure, the cost allocation (transfer pricing) gets more robust and thereby improves the stability of the network. For this reason, the third paper contributes to the second challenge to build more robust models (Fink et al. 2014).

2 Vehicle routing with private fleet, multiple common carriers offering volume discounts, and rental options

Authors: Christian Gahm, Christian Brabänder, Axel Tuma

This chapter is the peer-reviewed and published version of Gahm, Christian, Christian Brabänder, and Axel Tuma "Vehicle routing with private fleet, multiple common carriers offering volume discounts, and rental options.", in: Transportation Research Part E: Logistics and Transportation Review, 2017, Vol. 97: pp. 192-216, DOI: 10.1016/j.tre.2016.10.010.

Abstract

The problem addressed in this paper extends the vehicle routing problem with private fleet and common carriers by three aspects: two types of rental options, a cost function considering volumes and distances, and volume discounts offered by the common carriers. For its solution, we present a mixed integer program and three heuristics based on Variable Neighborhood Search. The computational analysis demonstrates the suitability of these heuristics and the positive effects of two newly introduced mechanisms. Analyzing the interdependencies between available outsourcing options and economic benefits, it shows that a subset of options is sufficient to reduce costs remarkable.

Keywords

Vehicle routing, private vehicles, common carrier, rental options, volume discounts, variable neighborhood search

2.1 Introduction

In this article, a comprehensive extension to the (capacitated) vehicle routing problem with private fleet and common carriers (VRPPC) is presented. The VRPPC, as discussed over the last ten years, tackles the problem of delivering products from a single central depot (e. g., shipping company) to customer locations. This task is accomplished either by the company's privately owned homogeneous or heterogeneous vehicle fleet (self-fulfillment) or by employing external common carriers (subcontracting), i.e., less than truckload (LTL) carriers. This standard VRPPC consists of a selection decision combined with a clustering decision and a routing decision. The first decision is to select one of the two delivery modes for each customer to be served; the second decision comprises the standard vehicle routing problem for the private fleet serving the assigned customers. Concerning the VRPPC, each customer must be served by exactly one vehicle of the limited private fleet or by exactly one external carrier (no split-delivery), every route of the private vehicles start and end at the depot, and vehicles of the private fleet have a specific capacity and perform at most one route per day. The objective is to minimize total delivery costs to serve all customers. Regarding the two delivery modes self-fulfillment and subcontracting, it is assumed that full truckload (FTL) deliveries executed by own vehicles are always cheaper than other delivery modes. Nevertheless, great saving opportunities are in outsourcing LTL deliveries to external carriers.

In this paper, we extend the VRPPC by three aspects. First, we consider two exclusive rental options as additional options for subcontracting: one with a rental fee charged on a route basis (mileage) and the second with a rental fee charged on a daily basis (see e. g., Krajewska and Kopfer 2009). Both rental options provide the opportunity to reduce total delivery costs by increasing the FTL delivery volume. The latter rental option is identical to the consideration of (full) truckload carriers that account for fixed cost per load up to a given capacity (see e. g., Rieksts and Ventura 2008 or Toptal and Bingöl 2011). Second, a more realistic concave freight function based on volumes and distances is integrated to determine the costs of LTL carriers (see e. g., Krajewska and Kopfer 2009). The third aspect is the consideration of volume discounts offered by (some) LTL carriers. Generally, volume discounts are a financial incentive used to foster demand and can be characterized as all-unit quantity discounts or incremental quantity discounts (Wilcox et al. 1987 and Weng 1995). From the view of a common carrier, these discounts provide potential to realize scale effects by freight consolidation in the short term (especially on the inbound route to a distribution center - see Nguyen et al. 2014; see Campbell 1990, on the effects of freight consolidation). In this context, Nguyen et al. (2014) discuss the increase of a customer's order size fostered by volume discounts offered to consolidate delivery orders. Discounts could also be used by LTL carriers to achieve a higher customer density, which leads to decrease in total delivery costs (see, Sun et al. 2015 on this relation). Therefore, the external carrier's operational costs per unit are likely to decrease and these savings could partially be passed to the shipper (to achieve a competitive advantage). Consequently, the shipper's delivery costs will also decrease. This and the previous aspect leads to a heterogeneous set of common carriers defining individually parametrized cost functions. Therefore, the new freight function for common carriers needs to be concave to represent decreasing freight rates depending on volumes and distances and needs to enable carrier-dependent discounts.

Altogether, these aspects lead to a new delivery planning approach for shipping companies addressed in this manner for the first time. We name this new approach the vehicle routing problem with private fleet, multiple common carriers offering volume discounts, and rental options (VRPPCdR). To be able to solve VRPPCdRs of virtually any problem size, new and enhanced solution methods based on the principles of Variable Neighborhood Search (VNS) are proposed. First enhancement is an explicit shaking mechanism for solution perturbation to support the exploration of the solution space. Second, a distance proportionate selection mechanism is introduced in order to improve the effectiveness of the local search procedures.

The Structure of the paper is as follows. The literature review in section 2.2 examines similar problems in detail and shows the relevance of the new planning approach. The VRPPCdR is formally described in section 2.3 and a mixed integer program (MIP) is presented there. The different solution methods and the introduced enhancements are described in section 2.4. The computational analysis in section 2.5 shows that the solution methods are suitable to solve the VRPPC and particularly the VRPPCdR. The analysis also shows that the new planning approach is able to reduce delivery costs remarkable and provides managerial insights on the effects of different subcontracting scenarios. Finally, conclusions and potential further research topics are described in section 2.6.

2.2 Literature review

Basically, two main research streams address operational transportation planning problems with subcontracting: one investigates the planning problem from the perspective of a freight-forwarding company; the other one investigates the perspective of a shipping company. The main difference between both streams lies in the basic planning problem: freight-forwarding companies have to solve a pick-up and delivery problem (PDP), whereas shipping companies owning a private fleet have to solve a capacitated vehicle routing problem (VRP). Of course, several variants of each of these basic problems are addressed in the literature. Another difference between a shipper and a forwarder is in the objective: While shippers aim to lower their total delivery costs, freight forwarders aim to acquire high volumes in order to generate revenues and lower costs per unit by consolidated full truck loads.

Since this paper takes the shipper's perspective, only a brief review of the relevant literature concerning the freight-forwarding perspective is given by selected papers. The paper of Krajewska and Kopfer (2009) is one of the first that combines several outsourcing options in an integrated manner. The authors enhance the underlying pick-up and delivery problem with time windows (PDPTW) by external carriers and the two exclusive rental options, described above. The authors called this problem the Integrated Transportation Planning Problem (ITPP; later also called Integrated Operational Transportation Planning - IOTP) and propose a tabu search heuristic extended by special types of moves for the different subcontracting options. In a similar context, Liu et al. (2010) address a task selection and routing problem in collaborative truckload transportation and solve the problem by a memetic algorithm. In contrast to the VRPPC, the authors include external delivery task during the shippers' distribution planning and the private fleet is of unlimited size. Based on the ITPP, Wang and Kopfer (2014) and Wang et al. (2014) formulate the Collaborative Transportation Planning (CTP) problem. Generally, collaborative planning can be seen as a joint decision-making process and CTP aims at the reallocation of requests among the partners in a horizontal cooperation. Accordingly, the main difference between the CTP and the ITPP is that CTP bases on an equal partnership, while in ITPP (and also the VRPPC) the players have a hierarchical relationship. Ziebuhr and Kopfer (2014) consider the IOTP from a forwarders perspective extended by compulsory requests, which are only permitted for self-fulfillment or premium subcontracting mode. Therefore, they use a formulation with common carriers and self-fulfillment and apply a large neighborhood search. Defryn et al. (2016) for example also address a vehicle routing problem in a collaborative environment: Based on the selective vehicle routing problem (SVRP), which can be interpreted as a VRPPC, different cost allocation methods for the SVRP in a collaborative environment are analyzed.

The first article that considers external carriers in the perspective of a shipping company originates from Ball et al. (1983). The authors investigate a fleet-size optimization problem, covering the option to outsource destinations to one external carrier. Klincewicz et al. (1990) and Hall and Racer (1995) also aim to determine an optimal fleet size when external carriers are available. When the private vehicle fleet contains only a single vehicle, the problem is similar to Prize Collecting Traveling Salesman Problem (PCTSP) without rewards for visiting nodes. Volgenant and Jonker (1987), Balas (1989), Bienstock et al. (1993), and Diaby and Ramesh (1995) tackled this problem (see also Balas 2007).

The VRPPC with a heterogeneous private fleet as described in section 1 is first examined by Chu (2005). Chu (2005) generally assumes that FTL deliveries are always cheaper to be delivered by own vehicles, whereas there might be saving opportunities in outsourcing LTL deliveries to an external carrier. In their paper, the costs of the external carrier are determined by a linear function of the Euclidean distance between the depot and the customer, but are independent of demand quantity. Chu (2005) proposes a 3-step heuristic: selection of customers to be served by the external carrier, initial solution construction by a modified savings algorithm (Clarke and Wright 1964), and solution improvement by a sequence of intra-route and inter-route exchanges. The Selection, Routing and Improvement heuristic (SRI) proposed by Bolduc et al. (2007) outperforms Chu’s heuristic by using two initial solutions and a λ-interchange procedure as first proposed by Osman (1993). Even better results for the VRPPC instances used by Chu (2005) and Bolduc et al. (2007) are reported in Bolduc et al. (2008). In contrast to Chu (2005) and Bolduc et al. (2007), who account for fixed costs independently of whether a private vehicle is used, Bolduc et al. (2008) account for these private vehicles’ fixed costs only if the corresponding vehicle is actually used (the authors also apply a linear function for the costs of the external carrier). Their Randomized construction Improvement-Perturbation heuristic (RIP) combines a descent strategy with two diversification strategies: a randomized savings construction phase and a perturbation mechanism. This RIP-heuristic also performs better (compared with SRI) when applied to instances originating from Christofides and Eilon (1969) and Golden et al. (1998). Côté and Potvin (2009) developed a tabu search heuristic (TS) that achieves better results than does the RIP heuristic when applied to the instances with a homogeneous private vehicle fleet (heterogeneous vehicles are not considered at all). With the extension of this tabu search heuristic by ejection chains (TS+), Potvin and Naud (2011) have further improved the results for homogeneous and heterogeneous private vehicle fleets. However, the computation times of the tabu search heuristics TS and TS+ are significantly greater than are those of the RIP heuristic.

Stenger et al. (2013b) extended the VRPPC by multiple depots and called their problem multi-depot vehicle routing problem with private fleet and common carriers (MDVRPPC). Their Adaptive Variable Neighborhood Search (AVNS) algorithm uses cyclic-exchange neighborhoods and incorporates an adaptive mechanism to bias the random shaking step. The AVNS is also evaluated by selected VRPPC instances (only instances with a homogeneous vehicle fleet are considered) and outperforms the RIP and TS algorithms. Furthermore, AVNS is almost equivalent to TS+ in terms of solution quality but requires significantly less computation time. Vidal et al. (2015) consider three different problems, each a particular variant of the vehicle routing problem with profits and customer selection. The authors propose a new large neighborhood search based on "exhaustive" solutions that are embedded within three heuristic frameworks (MS-LI, MS-ILS, and UGHS). Regarding the VRPPC, UGHS slightly outperforms AVNS in terms of solution quality but requires more than twice as much computation time.

All these papers focusing on the VRPPC consider solely the outsourcing to external LTL carriers. Only Kopfer and Wang (2009) consider all three outsourcing options. The authors extend the VRP to the Vehicle Routing and Forwarding Problem (VRFP) and their short evaluation based on very small test instances solved by the commercial solver CPLEX shows the economic benefit of sub-contracting in general.

There are several papers that consider volume discounts when planning the inbound transportation from suppliers (see e. g., Tersine and Barman 1991, Sheen and Tsao 2007, and Tsao and Lu 2012). To the best of our knowledge, only Stenger et al. (2013a) consider volume discounts in the outbound transportation (distribution) planning, when discussing routing with subcontracting. Here, Stenger et al. (2013a) extend the MDVRPPC of Stenger et al. (2013b) by quantity discounts and a minimum demand to be delivered by the private fleet. They call it the Prize-Collecting Multi-Depot Vehicle Routing Problem with Non-Linear costs (PCMDVRPNL) and propose a linear cost function and a non-linear stepwise cost function that depends on the vehicle capacity. This capacity is assumed to be identical for all vehicles of the private fleet and also for the vehicles of the LTL carrier. Both cost functions are linked with a maximal discount factor. Similar to Stenger et al. (2013a), we use a decreasing freight rate with a minimum freight rate but in contrast to them, we consider heterogeneous vehicles and heterogeneous common carriers having cost functions with individual parameters (e. g., discount factors).

Summarizing the literature review, there is no planning approach that simultaneously considers different rental options and volume discounts from the perspective of a shipping company.

2.3 Problem analysis

Before presenting a mixed integer program for the VRPPCdR, we specify the problem in detail and introduce notations in the following.

2.3.1 Problem specification and notation

The VRPPCdR considers four different transportation options for the delivery of products from the shipper's depot to customer locations. Basic planning tasks include, first, the selection of the transportation option for each customer to be served; second, the clustering of a subset of the set of customers to one available (own or rented) vehicle or to one available common carrier; and third, the routing of private and rented vehicles as necessary. The objective is to minimize total delivery (transportation) costs. Therefore, C represents the set of all vertices $h,i,j=0,...,n$ (representing customers and the depot, which is indexed by 0) and C^n, which represents the set of n customers where $n=|C^n|$. The distance between two vertices is defined by d_{ij} and the demand quantity of customer i by q_i.

The first transportation option is self-fulfillment by the heterogeneous private vehicle fleet of the shipping company. This set of m vehicles is denoted by V ($m=|V|$). The heterogeneous, limited capacity for each vehicle k is defined by cap_k and the variable cost rate per distance unit is defined by c_k. The route of each of these private vehicles must start and end at the shipper's depot. Although each of the own vehicles is associated with fixed costs c_k^{Fix}, we assume that these costs are not relevant in the short-term planning process but only when investigating different fleet sizes. This assumption is in line with Krajewska and Kopfer (2009) but opposes other approaches (e. g., Chu 2005 or Bolduc et al. 2007). Nevertheless, to compare, it would be easy to incorporate fixed costs as an additive term in the objective function.

To increase the transportation flexibility of the shipping company to fulfill FTL deliveries, we consider two options for exclusive vehicle rental. The subcontractor providing the first rental option charges a fee based on the distance of the route covered by one of his vehicles k. Therefore, a cost rate c_k^{Dist} per distance unit for each of the m' vehicles of set R' ($m'=|R'|$) is given. The fixed vehicle costs of this subcontractor are partially settled by the cost rate c_k^{Dist}. Therefore, $c_k^{Dist} > c_k$ strictly holds. To assert a minimum contribution to fixed costs, a minimum distance d_k^{Min} per rentable vehicle $k \in R'$ is defined by the subcontractor. Fees for the second rental option are accounted for on a daily basis. The subcontractor, who is paid on this daily basis, offers m'' vehicles from set R'' ($m''=|R''|$). These vehicles are charged for at cost rate c_k^{Day} if vehicle k is utilized in the resulting delivery plan. Here, the maximum utilization of vehicle k is limited by maximum distance d_k^{Max} per vehicle $k \in R''$. A maximum travel time could also be specified and converted to a maximum distance using standard travel speeds. All rentable vehicles also have a limited capacity cap_k and must start and end their routes at the shipper's depot. Because we assume that the company offering rental options is closely located to the shipper's depot, we neglect distances and travel times between the shipper's depot and these companies.

For all vehicles $k=1,...,m,...,m+m',...,m+m'+m''$ from the sets V, R', and R'', a subset of customers must be assigned (clustering), and a travel route must be determined (routing). As a result, those routes define the transportation costs either by the total traveled distance of the route ($k \in V \cup R'$) or by whether vehicle k is utilized at all ($k \in R''$).

The fourth transportation option is offered by a set E of $e=|E|$ common LTL carriers (logistics service providers or freight-forwarding companies). It is assumed that these carriers do not have any capacity limits but accept every subcontracted volume. Each of these carriers l defines individual parameters that are used to calculate the transportation costs for this type of subcontracting. In most publications, the cost for delivering a product to the customer only depends on the customer's locations and is represented by a linear function of the distance between the customer and depot (e. g., Chu 2005 or Bolduc et al. 2007). In contrast, Vahrenkamp (2011) states that a reasonable cost function must increase in both the quantity of delivered units and the distance from depot to customer location (as is common in public postal services). Krajewska and Kopfer (2009) also use this approach for the ITPP, defining a freight function by the weight of the cargo q_{ij}, the distance d_{ij} between two vertices i and j, a constant freight rate cfr, and the parameter $\lambda \in (0,1)$:

$$fr\left(d_{ij}, q_{ij}\right) = cfr \cdot \left(d_{ij} \cdot \left(q_{ij}\right)^{\lambda}\right)^{1-\lambda} \tag{1}$$

Transferring this freight function (1) to the VRPPCdR to determine the costs for serving customer $i \in C^n$ with one of the heterogeneous external carriers $l \in E$, the following (non-linear) cost function is used:

$$cf_{il} = tr_l \cdot \left(d_{0i} \cdot \left(q_i\right)^{\lambda_l}\right)^{1-\lambda_l} = tr_l \cdot \omega_{il} \tag{2}$$

In cost function (2), tr_l represents a constant tariff rate, and parameter $\lambda_l \in (0,1)$ defines the slope of the function (ω_{il} is introduced to increase readability in the following): For $\lambda_l = 1$, the cost function is a horizontal line at the value of tr_l. For $\lambda_l = 0$, the cost function is a linear function with slope tr_l and is independent of quantity q_i. Therefore, the usually used cost function (e. g., by Chu 2005 or Bolduc et al. 2008) is a special case of (2) in which $\lambda = 0$. The actual parameter values of cost function (2) and also of the volume discounts described in the following strongly depend on the goals (e. g., to acquire new customers) and the cost estimation of the common carrier (for the estimation of distribution costs see e. g., Turkensteen and Klose 2012 or Sun et al. 2015).

In addition to the extension of the VRPPC by rental options and cost function (2), we further consider all-unit volume discounts offered by LTL carriers. Here, we assume that the tariff rate tr_l itself is a function depending on the total delivery quantity outsourced to one of the LTL carriers $l \in E$. Therefore, the volume discount δ_l granted by carrier l is represented by a monotonically increasing function of the total quantity outsourced to carrier l (defined by the set O_l of customers assigned to carrier l). Instead of this function, also a piecewise linear function representing incremental freight discounts could be easily incorporated (like in Sheen and Tsao 2007, Tsao and Lu 2012, or Stenger et al. 2013a). In this paper, the discount is defined by a discount factor df_l.

$$\delta_l = df_l \cdot \sum_{i \in O_l} q_i \tag{3}$$

To prevent volume discounts increasing to a level above the undiscounted tariff rate tr_l, a minimal transportation rate tr_l^{Min} for every carrier l is used. Özkaya et al. (2010) discuss a similar approach to determine prices of LTL shipments (see also Stenger et al. 2013a). The following cost function integrates all of the previous considerations:

$$cf_{il} = \max\left\{tr_l - \left(df_l \cdot \sum_{i \in O_l} q_i\right),\, tr_l^{Min}\right\} \cdot \left(d_{0i} \cdot \left(q_i\right)^{\lambda_l}\right)^{1-\lambda_l} = \max\left\{tr_l - \delta_l,\, tr_l^{Min}\right\} \cdot \omega_{il} \qquad (4)$$

Unfortunately, the necessary maximum function (to guarantee a minimum transportation rate for the common carrier) leads to a non-linear cost function. This function will be linearized in the following to be able to provide a reasonably solvable mathematical formulation (the non-linearity of ω_{il} is not important in this context, because as a coefficient, it is applied independently of the selection and clustering decisions).

2.3.2 Mixed integer program

The objective function of the VRPPCdR consists of four parts, one for each of the transportation options. This approach has been chosen because of the multiple carriers with individual cost parameters and rental options. Thus, we do not expand the formulation for the heterogeneous VRP (HVRP) as proposed by Bolduc et al. (2008), but follow the modeling approach presented in Krajewska and Kopfer (2009).

Within the mathematical formulation, the following decision variables are used:

$$X_{ijk} = \begin{cases} 1 & \text{if vehicle } k \in V \cup R' \cup R'' \text{ visits vertex } j \text{ immediately after vertex } i \\ 0 & \text{otherwise} \end{cases}$$

$$X_k = \begin{cases} 1 & \text{if vehicle } k \in R'' \text{ is utilized (leaves the depot)} \\ 0 & \text{otherwise} \end{cases}$$

$$Y_{ik} = \begin{cases} 1 & \text{if vehicle } k \in V \cup R' \cup R'' \text{ visits vertex } i \\ 0 & \text{otherwise} \end{cases}$$

$$Z_{il} = \begin{cases} 1 & \text{if customer } i \text{ is assigned to carrier } l \in E \\ 0 & \text{otherwise} \end{cases}$$

Based on these variables and the redefinition of (3) to $\delta_l = df_l \sum_{j \in C^n} Z_{jl} \cdot q_j$, the objective function can be defined as follows:

$$\begin{aligned} \min \; & \sum_{k \in V} c_k \sum_{i \in C} \sum_{j \in C, i \neq j} X_{ijk} \cdot d_{ij} \\ & + \sum_{k \in R'} c_k^{Dist} \sum_{i \in C} \sum_{j \in C, i \neq j} X_{ijk} \cdot d_{ij} + \sum_{k \in R''} X_k \cdot c_k^{Day} \\ & + \sum_{l \in E} \sum_{i \in C^n} Z_{il} \cdot \max\left\{tr_l - \left(df_l \cdot \sum_{j \in C^n} Z_{jl} \cdot q_j\right), tr_l^{Min}\right\} \cdot \left(d_{0i} \cdot \left(q_i\right)^{\lambda_l}\right)^{1-\lambda_l} \end{aligned} \qquad (5)$$

Because the maximum operator of the cost function for the external carriers is not linear, this part of the objective function (5) is linearized to allow use of linear (standard) solution techniques to solve at least small VRPPCdR instances. This linearization (illustrated in detail in Appendix A of this chapter) leads to the subsequent MIP for the VRPPCdR. The linear objective function is stated in formula (6), and the constraints are given in formulas (7) to (21). Formulas (22) to (25) define the domain of the decision variables.

$$\begin{aligned}\min \; & \sum_{k\in V} c_k \sum_{i\in C} \sum_{j\in C, i\neq j} X_{ijk}\cdot d_{ij} \\ & + \sum_{k\in R'} c_k^{Dist} \sum_{i\in C}\sum_{j\in C, i\neq j} X_{ijk}\cdot d_{ij} + \sum_{k\in R''} X_k \cdot c_k^{Day} \\ & + \left(\sum_{l\in E} tr_l \sum_{i\in C^n} P_{il}\cdot \omega_{il}\right) - \left(\sum_{l\in E} df_l \sum_{i\in C^n}\sum_{j\in C^n} Q_{ijl}\cdot\omega_{il}\cdot q_j\right) + \left(\sum_{l\in E} tr_l^{Min} \sum_{i\in C^n}\left(Z_{il}-P_{il}\right)\cdot\omega_{il}\right)\end{aligned} \tag{6}$$

subject to

$$\sum_{i\in C^n} Y_{ik}\cdot q_i \leq cap_k \qquad \forall k\in V\cup R'\cup R'' \tag{7}$$

$$\left(\sum_{k\in V\cup R'\cup R''} Y_{ik}\right) + \left(\sum_{l\in E} Z_{il}\right) = 1 \qquad \forall i\in C^n \tag{8}$$

$$\sum_{i\in C} X_{ijk} = Y_{jk} \qquad \forall k\in V\cup R'\cup R'',\; j\in C \tag{9}$$

$$\sum_{j\in C} X_{ijk} = Y_{ik} \qquad \forall k\in V\cup R'\cup R'',\; i\in C \tag{10}$$

$$Y_{0k} \leq 1 \qquad \forall k\in V\cup R'\cup R'' \tag{11}$$

$$\sum_{j\in C^n} X_{0jk} = X_k \qquad \forall k\in R'' \tag{12}$$

$$\sum_{i\in C}\sum_{j\in C, i\neq j} d_{ij}\cdot X_{ijk} \leq d_k^{Max} \qquad \forall k\in R'' \tag{13}$$

$$\sum_{i\in C}\sum_{j\in C, i\neq j} d_{ij}\cdot X_{ijk} \geq d_k^{Min}\cdot Y_{0k} \qquad \forall k\in R' \tag{14}$$

$$U_i \leq \left(\sum_{k\in V\cup R'\cup R''} Y_{ik}\cdot cap_k\right) + \left(\sum_{l\in E} Z_{il}\cdot q_i\right) \qquad \forall i\in C^n \tag{15}$$

$$\begin{aligned} U_j \geq U_i + q_i + & \left(\sum_{k\in V\cup R'\cup R''} X_{ijk}\cdot cap_k\right) \\ & - \left(\sum_{k\in V\cup R'\cup R''} Y_{ik}\cdot cap_k\right) - \left(\sum_{l\in E} Z_{il}\cdot q_i\right) \end{aligned} \qquad \forall i,j\in C^n \text{ with } i\neq j \tag{16}$$

$$U_i \geq q_i \qquad \forall i\in C^n \tag{17}$$

$$Z_{il} - P_{il} \geq 0 \qquad \forall i\in C^n, l\in E \tag{18}$$

$$Z_{il} + Z_{jl} + P_{il} - 3Q_{ijl} \geq 0 \qquad \forall i, j \in C^n, l \in E \tag{19}$$

$$\left(tr_l \cdot Z_{il} - \delta_l\right) - tr_l^{Min} \leq P_{il} \cdot M \qquad \forall i \in C^n, l \in E \tag{20}$$

$$\left(tr_l - \delta_l\right) - tr_l^{Min} \geq \left(P_{il} - 1\right) \cdot M \qquad \forall i \in C^n, l \in E \tag{21}$$

$$X_{ijk}, Y_{ik} \in \{0,1\} \qquad \forall k \in V \cup R' \cup R'', i, j \in C \tag{22}$$

$$X_k \in \{0,1\} \qquad \forall k \in R'' \tag{23}$$

$$P_{il}, Q_{ijl} \in \{0,1\} \qquad \forall i \in C^n, l \in E \tag{24}$$

$$U_i \in \mathbb{Z}_{\geq 0} \qquad \forall i \in C^n \tag{25}$$

The capacity constraints (7) are defined for all own and rented vehicles. Constraint set (8) represents the single-delivery constraint. Constraint sets (9) and (10) ensure that the same vehicle enters and leaves a customer exactly once, and constraint set (11) ensures that all routed vehicles start from the depot. Constraint set (12) establishes the relationship between the "renting variable" X_k and the "routing variable" X_{ijk}. The maximal route length for vehicles rented on a daily basis is obeyed by (13), whereas (14) ensures the minimal distance of routes performed by vehicles charged per distance unit. The next constraint sets (15), (16), and (17) are sub-tour elimination constraints. Those are Miller-Tucker-Zemlin (MTZ) constraints (Miller et al. 1960) and are modified from the HVRP$_3$ formulation presented in Yaman (2006). Note, that we also tested formulations using the sub-tour elimination constraints as first proposed in Dantzig et al. (1954), but we found that for the VRPPCdR, the MTZ constraints provide a higher solution quality in shorter computation times. Constraint sets (18) to (21) are used for the linearization of volume discounts. Constraint set (18) allows the auxiliary variable P_{il} only being equal to 1 if $Z_{il} = 1$, but permits $P_{il} = 0$. The constraint set (19) works similarly for Q_{ijl}: All three variables Z_{il}, Z_{jl}, and P_{il} must be equal to 1 to permit $Q_{ijl} = 1$. Due to the minimization objective, the constraints defined by (20) and (21) prohibit $P_{il} = 1$ if $tr_l - \delta_l < tr_l^{Min}$, and vice versa. Here, the character M represents a sufficiently large number. For the formulation at hand, M can be approximated by $M \geq \max_{l \in E}\left(tr_l\right)$.

2.4 Solution Method

Because the VRPPCdR generalizes the NP-hard VRP and VRPPC (see Lenstra and Rinnooy Kan 1981 and Stenger et al. 2013b), it is also NP-hard. Thus, we develop three heuristic solution methods able to solve virtually any size of problem instances. Concerning the problem at hand, we especially focus on the development of very efficient heuristics, i.e., heuristics that are very fast in terms of computation time and achieve a sufficient good solution quality.

Starting with an initial solution calculated by a problem-specific construction heuristic, all three proposed heuristics are based on the principles of Variable Neighborhood Search (VNS; first described in Mladenović and Hansen 1997). VNS is applied because several papers show the suitability of VNS to solve VRPs and closely related problems (e. g., Hertz and Mittaz 2001, Bräysy 2003, Polacek et al. 2004, Kytöjoki et al. 2007, and Stenger et al. 2013b). Generally, VNS is a metaheuristic embedding a local search mechanism within a framework that uses multiple, varying neighborhood structures to exploit local optima and to escape from them (Hansen et al. 2010). These neighborhood structures are also used to define a solution perturbation mechanism ("shaking").

In addition to this well-known features of VNS, a completely new enhancing mechanism for VNS-based heuristics is introduced: ''explicit shaking". This enhancement is meant to explicitly interrupt the search procedure in already exhaustively searched areas of the solution space and thus, to direct the search process to other areas. The actual implementation of this explicit shaking mechanism is of course problem specific but the mechanism itself could be transferred to any VNS-based solution method. Furthermore, we introduce a distance proportionate vehicle selection mechanism to increase the likelihood of the local search for improvements.

All components of the proposed solution methods for the VRPPCdR, the problem-specific construction heuristic, the (new) VNS variants, the neighborhood structures and local search procedures, and the new explicit shaking mechanism are described in detail in the following sections.

2.4.1 Initial construction heuristic

The proposed construction heuristic for the initialization is a problem specific adaption of the procedure described in Bolduc et al. (2007). It integrates the two rental options and is able to handle multiple external carriers with individual parameters. Furthermore, we have modified the used savings heuristic (Clarke and Wright 1964) to the specific requirements defined by the two rental options. Additionally, the route construction procedure is terminated before the load of a vehicle k reaches 100% of its capacity (cap_k). The reason for this is that the succeeding solution improvement heuristic performs better if there is a certain amount of remaining capacity in each vehicle (also stated by Kytöjoki et al. 2007). This initial vehicle utilization is controlled by the parameter $\sigma \in \mathbb{R}$ $(0 < \sigma \leq 1)$. σ can be considered the percentage initial maximum utilization of every vehicle of the private fleet. The pseudocode Listing 1 illustrates the main course of action of the construction heuristic.

The heuristic starts with the calculation of the total customer demand (Listing 1-1) and the total capacity provided by the vehicles of the private fleet and all rentable vehicles (Listing 1-2; adjusted by σ). Here, we use the total capacity of the private fleet and all rentable vehicles because we follow the assumption that (full) truckloads lead to lower costs than LTL subcontracting for individual shipping.

Then, the set of customers C^n is sorted in ascending order by their mean external delivery costs (Listing 1-3). Because not only one external carrier is available but multiple ones, mean external delivery costs c_i^{Mean} are calculated for each customer *i* (naturally, volume discounts are not considered here):

$$c_i^{Mean} = \frac{1}{e} \cdot \sum_{l \in E} cf_{il} \tag{26}$$

Based on this ordering, a separation index *h* is calculated (Listing 1-4; see Bolduc et al. 2007) and C^n is divided into the sets $C^{V,R}$ and C^E (Listing 1-5) such that C^E contains the customers with the lowest mean external delivery costs and $C^{V,R}$ contains at most as many customers as can be served with respect to the adjusted capacity of the available vehicles in V, R', and R''.

1: td ← *getTotalDemand* (C^n);

2: tc ← *getTotalCapacity* (V, R', R'', σ);

3: customerList ← *sortCustomersByAscendingMeanExternalCarrierCosts* (C^n, E);

4: h ← *calculateSeparationIndex* (customerList, td, tc);

5: $C^{V,R}$, C^E ← *seperateCustomers* (customerList, h);

6: routes, remainingCustomerList ← *executeModifiedSavings* ($C^{V,R}$, V, R', R'', σ);

7: *merge*(C^E, remainingCustomerList);

8: *applyLKH-2*(routes);

9: *addExternalCarrierRoutes*(routes, *e*);

10: *assignRemainingCustomers*(routes, C^E);

Listing 1: Main course of action of the construction heuristic

The following modified sequential savings heuristic (Listing 1-6) constructs route by route as the original savings heuristic (Clarke and Wright 1964) but only as many routes as vehicles (including both the private fleet and the rentable vehicles) are available. The heuristic starts with the largest unused (i.e., no route is assigned) vehicle k and constructs feasible routes with respect to the adjusted capacity ($cap_k \cdot \sigma$) and, where necessary, with respect to minimum distance d_k^{Min} (if $k \in R'$) or to maximum distance d_k^{Max} (if $k \in R''$). Once all routes are constructed, the feasible route with the largest assigned demand quantity is assigned to k, all served customers are removed from $C^{V,R}$, and k is marked as in use. This procedure is repeated until either all customers are assigned to routes or no vehicle from V, R', or R'' is unused. Potentially unassigned customers are added to C^E (Listing 1-7).

Having all routes constructed and assigned, these routes are optimized (Listing 1-8) by Helsgaun's LKH-2 implementation (see Helsgaun 2000; Helsgaun 2009; available at http://www.akira.ruc.dk/~keld/research/LKH/) of the Lin-Kernighan algorithm (Lin and Kernighan 1973). This heuristic is used because it is one of the most powerful heuristics to solve Traveling Salesman Problems (Johnson and McGeoch 2007; Karapetyan and Gutin 2011) and, thus, for intra-route optimization.

To model subcontracting with external carriers, we use the concept of “virtual vehicles” (as, e. g., proposed by Bolduc et al. 2007; Stenger et al. 2013b), and *e* virtual routes are added to the list of “real” routes (Listing 1-9). This concept of virtual vehicles is used by all proposed solution methods. Finally, all customers from C^E are assigned to the external carrier with the lowest cost (calculated by cost function (2)) for serving this customer (Listing 1-10). Because all constraints are considered throughout the construction of the solution, the initial solution is always feasible. This feasibility of a solution is also ensured by all subsequently described solution transformations, and only feasible solutions are considered by the solution methods.

2.4.2 Main course of action of the VNS-variants

To solve the VRPPCdR, we propose a basic VNS variant as described in Hansen et al. (2010), a randomized Variable Neighborhood Descent (RVND) variant, and a VNS variant (named VNSac) that uses an acceptance criterion such as the Simulated Annealing (SA) metaheuristic. The RVND variant is inspired by the RandVND metaheuristic proposed in Gahm et al. (2014) and extends the basic Variable Neighborhood Descent (VND; Hansen et al. 2010) method by a random component. For all variants proposed here, we randomly select two vehicles (routes) before starting the neighborhood search. Therefore, we do not investigate all two-route combinations as was done by Kytöjoki et al. (2007) but only two randomly selected ones. Due to the random selection procedure, the RVND-variant is no longer deterministic but stochastic. The VNSac variant is developed because Hemmelmayr et al. (2009) reported a superior behavior of their SA-based VNS approach compared with their Skewed VNS and their VNS approach incorporating an acceptance criterion based on the Threshold Accepting concept. The main course of action for all three variants is described by the pseudocode in Listing 2.

The course starts with the initialization of the best known solution s^* and the incumbent solution $\tilde{s}$ by the construction heuristic. After this initialization, the main control loop is traversed for a maximum number of iterations i^{MAX} (Listing 2-2). The maximum number of iterations can be used by the decision-maker to balance the trade-off between solution quality and computation time.

1: $s^* \leftarrow \tilde{s} \leftarrow$ *getInitialSolution* ();
2: **loop1**: execute i^{MAX} iterations
3: *vehicle1*, *vehicle2* ← *getTwoRandomVehicles* ($\tilde{s}$, ζ);
4: **loop2**: iterate through neighborhood structures
5: **if** (**not** *RVNDisActive*) **then**
6: *applyShaking* (*vehicle1, vehicle2*, *getRandomOperator*());
7: **end if**;
8: s' ← *executeLocalSearch* (*vehicle1*, *vehicle2, additional parameters*);
9: **if** (*costs*(s') < *costs*(s^*)) **then** // new ever best solution
10: $s^* \leftarrow \tilde{s} \leftarrow s'$; **exit loop2**; // stop iterating through the neighborhood structures

11: **else**
12: **if** (*costs*(s') < *costs*($\tilde{s}$) // improvement of the incumbent solution
13: **or** (*VNSacIsActive* **and** *isDegradationAcceptable* (s' , $\tilde{s}$, *T*)) **then**
14: $\tilde{s} \leftarrow s'$; **exit loop2**;
15: **end if**;
16: **end if**;
17: **end loop2**;
18: **if** ($\kappa > 0$ **and** i^{WI} is reached) **then**
19: *executeExplicitShaking* ($\tilde{s}$, *vehicle1*, *vehicle2*);
20: **end if**;
21: **if** ($\kappa > 0$ **and** i^{ES-WI} is reached) **then**
22: $\tilde{s} \leftarrow s^*$; // reset explicit shaking
23: **end if**;
24: **if** (*VNSacIsActive*) **then** *T* ← *updateTemperature*(*T*, τ);
25: **end loop1**;

Listing 2: Main course of action of the VNS-variants

Regardless of the specified variant, each iteration starts with the random selection of two different vehicles/routes (Listing 2-3). Here, to increase the likelihood for improving moves within the subsequent local search, a new distance proportionate vehicle selection mechanism is introduced (see Section 2.4.4 for details).

The inner control loop (Listing 2-4) iterates through the defined neighborhood structures ns_1 to ns_5 as described in detail in section 2.4.3 (see Table 2.1). Then, for the two variants VNS and VNSac, a randomly selected operator (from the currently active neighborhood structure) performs a shaking step for the two randomly selected vehicles (Listing 2-6). Because the RVND variant bases on VND, no shaking step is executed (see Hansen et al. 2010).

The current ("shaked") solution is then improved by a specified local search procedure and returns the new solution s' (Listing 2-8; the local search procedure is described in the following section).

When the costs of solution s' (calculated by cost function (6)) are less than that of the best known solution s^*, solution s' becomes s^* and also the new incumbent solution $\tilde{s}$ and iterating through the neighborhood structures is terminated (Listing 2-9…10). If only the costs of incumbent solution $\tilde{s}$ are improved by s' (Listing 2-12) or the VNSac variant is active and the cost degradation is acceptable (Listing 2-13), then the new solution s' becomes $\tilde{s}$ and iterating through the neighborhood structures is also terminated (Listing 2-14). The acceptance probability of non-improving solutions is defined by $exp\left(-\left(costs(s'') - costs(\tilde{s})\right)/\mathrm{T}\right)$, where T represents the (current) temperature. In any other case, iterating through the neighborhood structure is continued.

If the explicit shaking mechanism is activated at all ($\kappa > 0$) and a specified number of iterations without improvement i^{WI} is reached, an "explicit shaking" mechanism as described in section 2.4.5 is executed (Listing 2-18…20). To avoid the exploration of unpromising parts of the solution space when explicit shaking is activated, a mechanism is available to reset the incumbent solution $\tilde{s}$ to the best known solution s^* if no improvement of s^* can be achieved within i^{ES-WI} iterations (Listing 2-21…23).

The last step within loop1 is the update of the temperature for the VNSac variant (Listing 2-24). To update the temperature, we use a linear cooling scheme based on cooling factor τ : $T_{t+1} = T_t \cdot \tau$. Different cooling schemes (exponential cooling and step functions) have also been considered, but linear cooling with $\tau = 0.3$ provided the best results.

2.4.3 Neighborhood structures and local search procedure

The most important aspects when applying VNS to a specific problem are the definition and number of appropriate neighborhood structures (Listing 2 loop2), the order in which these neighborhoods are investigated, the strategy that is used for changing neighborhoods, and the local search procedure (Hansen and Mladenović 2001).

Neighborhood structures and also the local search procedures are based on transformation rules, which define the transformation of a solution to obtain another solution. For the VRPPCdR, we use increasingly large neighborhood structures with changing local search mechanisms based on λ-interchanges (Osman 1993). Each neighborhood structure is defined by a set of operators that define the λ-interchanges and the extraction/insertion (e/i) mechanism to use. According to Osman (1993), shifting or transferring one customer from one route to another is defined by the (1, 0) and (0, 1) operator, respectively. Specifically, the operator (1, 0) means that a single customer is extracted from one route and inserted into another route. Shifting two customers is defined by (2, 0) and (0, 2), shifting three by (3, 0) and (0, 3), and so on. Interchanging or swapping customers between two routes is defined by operators such as (1, 1) and (2, 2). In contrast to Stenger et al. (2013b), we only consider two routes for inter-route optimization because this approach is also used by the very performant VNS metaheuristic presented in Kytöjoki et al. (2007). To improve the efficiency of the inter-route optimization, we introduce a new mechanism to randomly select these two routes. This mechanism is described in section 2.4.4.

To extract and insert customers from and into a route, we use two local-search mechanisms: "sequence-e/i" and "single-e/i". To transfer promising sequences between two routes, the sequence e/i-mechanism extracts customers in sequence and inserts this sequence into the other route, preserving the sequence. In contrast, the single e/i-mechanism extracts randomly chosen customers and inserts them individually. With this mechanism, the sequence of both routes is broken up. To insert either the complete sequence of customers or all customers individually, the position with the lowest cost is calculated for insertion. Of course, if customers are going to be transferred to virtual routes, no insertion position is calculated but only the costs. Due to the heterogeneous vehicle fleet considered by the VRPPCdR, it is not possible to use the very fast cost estimation procedure as described in Bolduc et al. (2008). After applying one of the operators, the routes are tested for cost improvements and for feasibility and, if they are feasible, an intra-route optimization step is performed by 2-opt exchanges (Lin 1965). We use 2-opt exchanges because tests have shown that more-sophisticated methods for intra-route optimization such as the outstanding LKH-2 heuristic have only marginal positive effects on the solution quality but significantly increase computation time.

Table 2.1 Neighborhood structures

Neighborhood structure	e/i-mechanism	λ-operator sets
ns_1	single-insertion	{(1,0), (0,1), (1,1)}
ns_2	sequence-insertion	{(2,0), (0,2), (2,1), (1,2), (2,2), (3,0), (0,3), (3,1), (1,3), (3,2), (2,3), (3,3)}
ns_3	single-insertion	
ns_4	sequence-insertion	{(4,0), (0,4), (4,1), (1,4), (4,2), (2,4), (4,3), (3,4), (4,4), (5,0), (0,5), (5,1), (1,5), (5,2), (2,5), (5,3), (3,5), (5,4), (4,5), (5,5)}
ns_5	single-insertion	

Based on preliminary tests, we also identified the sequence of five neighborhood structures ns_1 to ns_5 listed in Table 2.1 as most appropriate concerning solution quality and computation time.

Input data and parameters for the local search procedure applied to each neighborhood structure are two routes, one operator, the e/i-mechanism and the improvement strategy. The improvement strategy parameter ψ defines the manner in which solutions will be selected to become the incumbent solution. The strategies "best improvement/steepest decent (BI)" and "first improvement/first descent (FI)" are used (for details see, e. g., Hansen et al. 2010). The BI-strategy evaluates all possible solutions defined by one neighborhood structure and returns the best of these solutions. In contrast, using the FI-strategy, the evaluation of a neighborhood structure stops as soon as some improvement is detected. Obviously, this parameter influences the number of neighboring solutions to be explored and thus, has a strong influence on the computation time. This influence on the computation time and on the solution quality will be analyzed in detail within the computational analysis in section 2.5.

2.4.4 Vehicle selection mechanism

Two mechanisms are used to select the two vehicles for inter-route optimization (local search). Parameter ζ controls their usage. The first mechanism (ζ = 0) randomly selects two vehicles from all available real and virtual vehicles. The second mechanism (ζ = 1) also selects vehicles randomly, however, if the first selected vehicle $v1$ is a real vehicle (in one of the sets V , R', or R''), then we apply a distance proportionate mechanism for the selection of the second vehicle $v2$. If the first vehicle represents a common carrier, the second one is randomly selected from all other available vehicles. In this case, all vehicles have a uniform probability to be selected for first or second vehicle.

The basic idea of the distance proportionate selection mechanism is to guide the selection of vehicles for inter-route optimization in a way that increases the probability of finding an improvement. Vehicle routes in close proximity to the first selected vehicle (route) have a greater probability for a successful inter-route improvement move than routes that are far away. To determine the proximity of two routes, the Euclidean distance of their center points cp_{v1} and cp_{v2} is used. The center point cp_v of a route (vehicle) is calculated by the centroid of the axis-aligned minimum bounding rectangle (AAMBR) of all customers on the route (

Fig. 2.1 illustrates the calculation of the x-coordinate and the y-coordinate of a center point). Of course, there are more (sophisticated) ways to calculate a center of a route (like the geometric median or solving a p-median problem). However, we decided for the AAMBR due to its low computational effort. For all common carriers, we use the mean distance of the real vehicles.

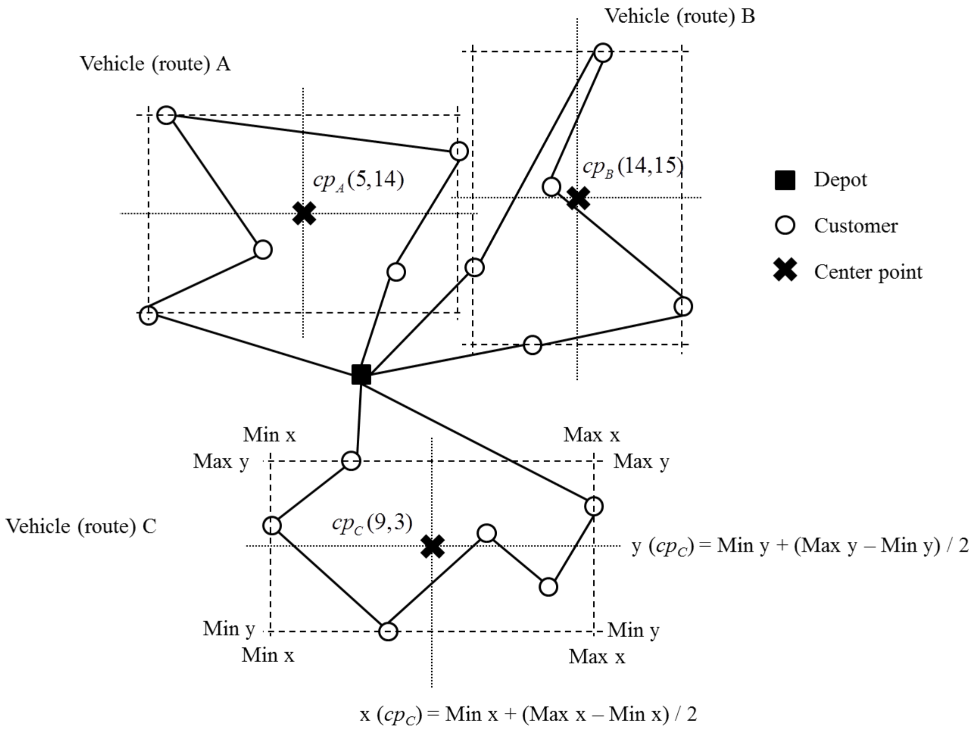


Fig. 2.1 Route center calculation with AAMBR

The intended effect can be best illustrated considering the vehicle routes in Fig. 2.1. Assuming that vehicle A with $cp_A(5,14)$ is selected as first vehicle, the probability that vehicle B with $cp_B(14,15)$ and $dist_{A,B} = 9.06$ is chosen as second vehicle should be higher than that of vehicle C with $cp_C(9,3)$ and $dist_{A,C} = 11.7$. Therefore, the selection is based on a reciprocal distance proportionate mechanism. This procedure is inspired from fitness proportional selection often used in Genetic Algorithms (see e. g., Grefenstette 2000).

The effect of the distance based vehicle selection is analyzed in section 2.5.2.

2.4.5 Explicit shaking mechanism

To support the exploration of the solution space and to avoid cycling in local optima, we use two additional mechanisms for solution perturbation. Because both mechanisms have a high intended perturbation effect, these rules are not used to define neighborhood structures but rather to define "explicit" shaking mechanisms (Listing 2-18…20).

The first of these very specific transformation rule is called "route swapping" (RS) because it swaps complete routes between two vehicles (similar to Krajewska and Kopfer 2009). In contrast to Krajewska and Kopfer (2009), we swap routes not only between own and rented vehicles but between all (real and virtual) vehicles. The second transformation rule "route dividing" (RD) has also a very large perturbation effect and is only used if route swapping is not possible due to capacity or distance constraints (see (7), (13), and (14)). The RD-rule takes a real vehicle's route, extracts all customers, and inserts each of these customers to the external carrier with the lowest costs for serving this customer. After applying RD, one vehicle has an empty route and thus great potential for shifting customers to this route.

In principle, these explicit shaking mechanisms are applied whenever it is not possible to improve the best known solution s^* for a certain number of iterations i^{WI} (see Listing 2-20…22). To avoid a too exhaustive, ineffective exploration of unpromising regions of the solution space after an explicit shaking step, we stop the search after a predefined maximum number of iterations (i^{ES-WI}) without improvement of s^*, reset to the best known solution thus far (s^*), and perform another explicit shaking step (see Listing 2-23…25). The usage of the explicit shaking mechanisms is controlled by the parameter $\kappa \in \mathbb{R}$ (with $0 \le \kappa < 1$), whereby $\kappa = 0$ has an additional meaning and avoids any explicit shaking. The additional parameter θ specifies whether the RD-transformation rule is used as an explicit shaking mechanism at all.

In the next section, the explicit shaking mechanisms RS and RD and the specific parameter settings for κ and θ are analyzed with respect to their influence on the performance of the three VNS-variants. Furthermore, this analysis covers the influence of the improvement strategy parameter ψ and the vehicle selection mechanism controlled by ζ.

2.5 Computational analysis

The computational analysis consists of two parts. In the first part, the performance of the proposed VNS-variants combined with different parameter settings is evaluated by the variants' solution quality and computation time. The second part assesses the economic benefits of considering multiple external carriers offering volume discounts and rental options. Furthermore, the particular circumstances of achieving these benefits are analyzed.

2.5.1 Problem instances and basic parameter settings

Before presenting the analysis, the used test instances and the basic parameter settings are described in this section. The first two sets of instances CHU and BOL contain the instances used in Chu (2005) and Bolduc et al. (2007). Unfortunately, we discovered some inconsistency using the objective values presented in these papers (we think caused by different rounding of the Euclidean distances); therefore, we calculated new objective values with the model described above and the commercial standard solver IBM CPLEX Optimizer 12.5 (the new reference values can be found in Appendix C, Table 2.14). To handle the rounding issue for these instances, we calculate the Euclidean distance $d_{ij} \in \mathbb{N}$ between two vertices *i* and *j* by their coordinates in advance by

$$d_{ij} = \left\lfloor \sqrt{\left(x_i - x_j\right)^2 + \left(y_i - y_j\right)^2} \right\rfloor.$$

The next four instance sets are based on the instances of Christofides and Eilon (1969) and Golden et al. (1998) and have been adapted by Bolduc et al. (2008) for the VRPPC. These four instance sets are named CE and G for instances with a homogeneous private vehicle fleet and CE-H and G-H for instances with a heterogeneous private vehicle fleet. The instances are available at http://www.mcbolduc.com/VRPPC/tests.htm. Reference values for instances in sets CE and G are taken from Stenger et al. (2013b); specifically, the objective values with their mean computation times as listed in "Table 4: Results on VRPPC Instances" in the paper of Stenger et al. (2013b) (but not the values in column "BKS" because computation times for these values are not given). For the instances in the sets CE-H and G-H, we use the objective values with their mean computation times as listed in "Table 8: Heterogeneous instances" in the paper of Potvin and Naud (2011).

For the VRPPCdR, we define seven new instance sets based on the BOL-instance set. We combine each of the basic scenarios BOL-1 to BOL-5 (basic data can be found in Appendix C, Table 2.15) with a downsized private fleet (vehicle data are listed in Table 2.2) and seven "carrier and rental option"-scenarios (Table 2.5). The data of the private vehicles, the capacities and the cost factors, are the same as used in Bolduc et al. (2007).

Table 2.2 Vehicle data of the private fleet ($k \in V$)

Identifier	cap_k	c_k	c_k^{Fix}
V1	40	1.5	60
V2	75	1.5	120
V3	110	1.5	150
V4	100	1.5	140
V5	4500	1.5	250
V6	4000	1.5	200

The parameters of rental options M1 to M3 (Table 2.3) and D1 and D2 (Table 2.4) are deduced from the original vehicles' parameters to obtain reasonable settings. Reasonable means that the cost factors for both rental options are specified according to the assumption that the cost rate per distance unit of the private fleet is always cheaper than the freight rate of any other delivery mode. In addition, this kind of parameter setting reflects the fact that shipping companies usually prefer the use of their own fleet before considering any type of subcontracting (not using the private vehicles is a kind of waste).

To get fitting parameter settings, the maximum capacity per rentable vehicle and the daily rate c_k^{Day} are calculated based on the corresponding maximum values of all vehicles of the private fleet *R* associated with the carrier and rental option scenario (the private vehicles data derive from Bolduc et al. 2007). Another factor that the cost factors c_k^{Dist} are greater than are those of the private vehicles is that they contain proportional fixed costs. The d_k^{Min} and d_k^{Max} parameters are estimated based on the length of the optimal routes reported in Bolduc et al. (2007).

Table 2.3 Rental options paid on a mileage basis ($k \in R'$)

Identifier	cap_k	c_k^{Dist}	d_k^{Min}
M1	$\lfloor \max\{cap_k : \forall k \in V\} \cdot 1.33 \rfloor$	2.50	60
M2	$\max\{cap_k : \forall k \in V\}$	2.25	60
M3	$\lfloor \max\{cap_k : \forall k \in V\} \cdot 0.5 \rfloor$	2.00	60

Table 2.4 Rental options paid on a daily basis ($k \in R''$)

Identifier	cap_k	c_k^{Day}	d_k^{Max}
D1	$\lfloor \max\{cap_k : \forall k \in V\} \cdot 1.33 \rfloor$	$\lfloor \max\{c_k^{Fix} : \forall k \in V\} \cdot 2.66 \rfloor$	120
D2	$\max\{cap_k : \forall k \in V\}$	$\max\{c_k^{Fix} : \forall k \in V\} \cdot 2$	120

The seven carrier and rental option scenarios S1 to S7 contain different combinations of rental options and external carriers; however, S7 contains all possible options (Table 2.5).

Parameters of the four considered external carriers are shown in Table 2.6. In turn, these parameters are specified according to the assumption that subcontracting should generally be more expensive than self-fulfillment.

Table 2.5 Carrier and rental option scenarios

Identifier	Rental options	Carriers
S1		C2, C3
S2		C1, C2, C3, C4
S3	M1, M3, D1, D2	C1
S4	M2, M3, M3, D2	C1
S5	M2, D2	C2, C3, C4
S6	M3, M3, D1, D2	C2, C3
S7	M1, M2, M3, D1, D2	C1, C2, C3, C4

The first two external carriers C1 and C2 offer no volume discounts and have a constant tariff rate tr_l. The tariff of C1 is independent of the delivered quantity, whereas C2 has a slope of $\lambda_l = 0.4$. Carriers C3 and C4 offer volume discounts and thus have a discount factor df_l and minimum tariff rate tr_l^{Min} specified (the effect of the specified volume discount parameters for the carriers C3 and C4 is illustrated in Fig. 2.2 in Appendix B).

Table 2.6 Carrier data ($l \in L$)

Identifier	tr_l	λ_l	df_l	tr_l^{Min}
C1	4	0.0	0	0
C2	7	0.4	0	0
C3	7	0.2	0.02	3
C4	5.5	0.2	0.015	4

Altogether, these specifications lead to 35 problem instances (Table 2.7) grouped into five sets N1 to N5 (with respect to their basic data based on the BOL-instance set), each comprising the seven carrier and rental option scenarios S1 to S7 listed in Table 2.5.

Table 2.7 New problem instances for the VRPPCdR

Instance identifier	Private fleet	Carrier and rental option scenarios	Basic data (locations and order quantities)
N1-01 … N1-07	V1	S1 – S7	BOL-1
N2-01 … N2-07	V2	S1 – S7	BOL-2
N3-01 … N3-07	V3, V4	S1 – S7	BOL-3
N4-01 … N4-07	V5	S1 – S7	BOL-4
N5-01 … N5-07	V5, V6	S1 – S7	BOL-5

The parameters listed in Table 2.8 are the result of comprehensive preliminary tests. With these parameter settings, best results can be achieved considering the trade-off between solution quality and computation time.

Table 2.8 Parameters for the three VNS-variants (based on preliminary tests)

Parameter	Description	Definition
ε	Maximum iteration factor	$\varepsilon = 1$
σ	Initial vehicle utilization	$\sigma = 0.85$
τ	Linear cooling factor	$\tau = 0.3$

i^{MAX}	Maximum number of total iterations	$i^{MAX} = \max\{200,\ n\cdot\left(m+m'+m''+e\right)\cdot\varepsilon)$
i^{WI}	Maximum number of iterations without new best known solution $\hat{s}$	$i^{WI} = \max\left\{30,\ n\cdot\left(m+m'+m''+e\right)\cdot\kappa\right\}$
i^{ES-WI}	Number of iterations without new best solution after explicit shaking	$i^{ES-WI} = i^{WI} - 1$

The effects on solution quality and computation time of parameters not listed in Table 2.8 are evaluated in detail in the next section.

2.5.2 Performance of the VNS-variants

The first part of this evaluation assesses the solution quality and computation times of the different VNS-variants at an aggregated level. Here, we consider all instance sets defined above (CHU, BOL, CE, G, CE-H, G-H, and N1 to N5) and compare the improvements in terms of computation time and solution quality (see Table 2.9 and Table 2.10). All (mean) computation times resulting from experiments using one of the proposed VNS variants are performed on a computer with an Intel® XEON® CPU E5 at 2.9 GHz. Because realistic reference computation times are only available for the instance sets CE, G, CE-H, and G-H, the relative computation time improvements ("CT^{Mean} imp. [%]") in Table 2.9 and Table 2.10 are only based on the results for these sets. To assess the relative improvement of the solution quality, we use the minimum objective values obtained by our proposed solution methods to calculate the relative improvement "OV^{Min} imp. [%]". Reference values for comparison are the best known objective value either obtained by IBM CPLEX Optimizer 12.5 (after 50 hours) or taken from literature. The minimum objective values result from ten experiments for each VNS-variant and parameter setting. The results for each instance of all instance sets are aggregated to a single value. Table 2.9 and Table 2.10 show these relative improvements for VNS, RVND, and VNSac with completely random vehicle selection (ζ = 0; Table 2.9) and with distance proportionate vehicle selection (ζ = 1; Table 2.10); and with respect to the parameters i^{MAX} (maximum number of iterations), ψ (improvement strategy), κ (explicit shaking factor; zero means no explicit shaking), and θ (RD-transformation rule, whereby T ("true") means that the RD-transformation rule is available). Bold values in the tables mark maximum improvements per column.

The numbers in Table 2.9 show that the FI-strategy is dominated by the BI-improvement strategy in terms of solution quality and that both explicit shaking mechanisms have a positive effect on the solution quality. The numbers also show that the RVND variant outperforms the other variants in terms of solution quality but at the expense of computation time. Regarding the maximum number of iterations, the effect of this parameter on the trade-off between solution quality and computation time comes obvious. Best setting with regard to solution quality is RVND with ζ = 0, i^{MAX} =1, ψ =BI, κ =0.1, and θ =T, which achieves a mean improvement of 0.35% (with a computation time improvement of 18.0%). Considering the best trade-off between solution quality and computation time, RVND with ζ = 0, i^{MAX} =0.5, ψ =BI, κ =0.1, and θ =F performs best and achieves a mean solution quality improvement of 0.15% and a mean computation time reduction by 68.6%.

Table 2.9 Relative improvements with ζ = 0 and with respect to the parameters i^{MAX}, ψ, κ, and θ

			VNS				RVND				VNSac			
		i^{MAX}	1		0.5		1		0.5		1		0.5	
ψ	κ	θ	OV^{Min} imp. [%]	CT^{Mean} imp. [%]	OV^{Min} imp. [%]	CT^{Mean} imp. [%]	OV^{Min} imp. [%]	CT^{Mean} imp. [%]	OV^{Min} imp. [%]	CT^{Mean} imp. [%]	OV^{Min} imp. [%]	CT^{Mean} imp. [%]	OV^{Min} imp. [%]	CT^{Mean} imp. [%]
FI	0	F	-2.61	82.5	-4.26	94.5	-1.47	16.7	-1.70	68.4	-2.08	85.2	-4.66	95.5
FI	0.7	F	-2.48	82.6	-3.21	94.2	-1.17	16.3	-1.69	68.5	-2.06	85.0	-4.36	95.8
FI	0.3	F	-2.07	83.6	-3.23	94.2	-0.58	20.2	-1.02	68.0	-2.06	85.2	-4.48	95.4
FI	0.1	F	-1.81	83.3	-3.42	94.5	-0.04	24.0	-0.46	67.8	-2.43	85.1	-4.49	95.7
FI	0.7	T	-2.26	83.4	-3.51	94.2	-1.09	16.6	-1.67	69.0	-2.55	84.3	-4.22	95.2
FI	0.3	T	-2.29	82.6	-3.15	94.8	-0.15	23.5	-1.43	69.3	-1.95	83.8	-4.74	95.8
FI	0.1	T	-1.81	82.8	-3.33	94.2	-0.19	35.7	-0.40	67.8	-2.15	84.0	-5.42	95.9
			-2.19	83.0	-3.45	94.4	-0.67	21.9	-1.19	68.4	-2.18	84.7	-4.62	95.6
BI	0	F	-0.63	48.6	-1.43	79.9	-1.36	9.8	-1.42	69.0	-0.31	51.6	-1.08	76.0
BI	0.7	F	-0.84	49.2	-0.98	79.9	-0.52	9.7	-0.12	69.2	-0.27	51.2	**-0.98**	76.5
BI	0.3	F	**-0.25**	50.0	-0.99	79.9	-0.03	12.1	0.14	69.6	-0.20	51.6	-1.16	75.7
BI	0.1	F	-0.27	48.3	-0.80	79.9	0.04	13.3	**0.15**	68.6	**0.00**	51.7	-1.14	75.6
BI	0.7	T	-0.80	50.5	-1.13	79.8	-0.62	11.0	-0.07	68.5	-0.25	49.9	-1.18	77.0
BI	0.3	T	-0.32	49.4	**-0.77**	79.8	-0.06	13.2	-0.05	69.7	-0.18	50.9	-1.01	76.2
BI	0.1	T	-0.39	48.3	-1.13	79.8	**0.35**	18.0	-0.02	73.6	-0.35	51.5	-1.41	75.7
			-0.50	49.2	-1.03	79.9	-0.31	12.4	-0.20	69.7	-0.22	51.2	-1.14	76.1

The effect of the distance proportionate vehicle selection mechanism is shown by Table 2.10. The numbers show that this mechanism is able to slightly improve the solution quality for almost all parameter settings with a similar computation time. For RVND with the first improvement strategy, even computation times can be reduced remarkable. For ζ = 1, the best solution method and parameter setting for i^{MAX} =1 is defined by RVND and ψ =BI, κ =0.3, and θ =F; here, a mean improvement of 0.19% (with a computation time improvement of 11.7%) is achieved. For ζ = 1 and i^{MAX} =0.5, RVND with ψ =BI, κ =0.3, and θ =T performs best and improves solution quality by 0.20% and reduces mean computation time by 69.8%.

Summarizing this first analysis, we conclude that the RVND-variant with ζ = 0, i^{MAX} =1, ψ =BI, κ =0.1, and θ =T (RVND^{SQ}) performs best in terms of solution quality and that RVND with ζ = 1, i^{MAX} =0.5, ψ =BI, κ =0.3, and θ =T (RVND^{TO}) performs best regarding the trade-off between solution quality and computation time. Of course, this trade-off always depends on the decision-maker's preferences.

Table 2.10 Relative improvements with ζ = 1 and with respect to the parameters i^{MAX}, ψ, κ, and θ

	VNS				RVND				VNSac			
i^{MAX}	1		0.5		1		0.5		1		0.5	
ψ κ θ	OV^{Min} imp. [%]	CT^{Mean} imp. [%]	OV^{Min} imp. [%]	CT^{Mean} imp. [%]	OV^{Min} imp. [%]	CT^{Mean} imp. [%]	OV^{Min} imp. [%]	CT^{Mean} imp. [%]	OV^{Min} imp. [%]	CT^{Mean} imp. [%]	OV^{Min} imp. [%]	CT^{Mean} imp. [%]
FI 0 F	-1.87	82.7	-3.95	94.8	-1.64	46.9	-1.65	67.9	-2.04	86.3	-4.74	95.6
FI 0.7 F	-2.53	82.6	-3.66	94.0	-1.10	45.5	-1.66	70.1	-2.10	86.1	-4.29	95.9
FI 0.3 F	-2.21	82.3	-3.25	94.1	-0.37	47.9	-0.65	68.5	-1.78	86.4	-4.51	95.8
FI 0.1 F	-1.40	82.2	-3.04	94.9	-0.02	49.4	-0.32	67.9	-1.73	86.4	-4.80	95.8
FI 0.7 T	-2.33	82.5	-3.20	94.1	-1.19	46.6	-1.50	68.2	-2.30	85.4	-4.79	95.6
FI 0.3 T	-1.83	82.9	-3.23	94.0	-0.36	49.9	-0.81	69.3	-2.12	86.6	-4.52	95.9
FI 0.1 T	-1.76	82.2	-3.28	94.0	-0.37	54.8	-0.29	67.6	-1.99	86.2	-4.93	95.8
	-1.99	82.5	-3.37	94.3	-0.72	48.7	-0.98	68.5	-2.01	86.2	-4.65	95.8
BI 0 F	-0.62	50.2	-1.45	79.8	-1.33	10.4	-1.39	69.2	-0.35	51.5	**-0.96**	78.2
BI 0.7 F	-0.45	47.9	-0.97	80.0	-0.30	10.1	0.18	69.3	-0.40	58.2	-1.30	78.2
BI 0.3 F	0.02	49.7	**-0.67**	80.2	**0.19**	11.7	0.16	69.7	-0.25	58.3	-1.02	78.3
BI 0.1 F	-0.36	48.0	-1.04	79.8	0.12	13.3	0.06	68.8	-0.24	51.0	-1.29	78.1
BI 0.7 T	-0.56	57.5	-1.02	77.0	-0.48	11.7	-0.03	69.4	-0.26	58.2	-0.98	78.2
BI 0.3 T	**0.16**	49.0	-0.99	79.9	0.08	12.9	**0.20**	69.8	-0.39	58.1	-1.06	78.3
BI 0.1 T	-0.13	48.4	-0.80	79.9	-0.04	18.4	0.10	69.7	**-0.21**	58.6	-1.27	78.2
	-0.28	50.1	-0.99	79.5	-0.25	12.7	-0.10	69.4	-0.30	56.3	-1.13	78.2

In the second part of this evaluation, we compare the objective values achieved by RVND^{SQ} and RVND^{TO} with the reference values of the instance sets CE, G, CE-H, and G-H to assess the solution quality in a more detailed way. Therefore, for each instance in these sets, Table 2.11 and Table 2.12 list the solution method (''SM''; for abbreviations, see Section 2.1) the best known objective value (''BKOV'') is achieved with and the mean computation time of the BKOV (''CT'') in seconds. For both parameter settings (RVND^{SQ} and RVND^{TO}), the minimum objective value (''OV'') from the ten calculated solutions, the corresponding relative improvement concerning BKOV and OV (''Imp.

OV"), and the mean relative computation time improvement ("Imp. CT"), are presented. In both tables, improvements of the BKOV are marked bold. Because no detailed computation times for MS-LI, MS-ILS, and UGHS are given in Vidal et al. (2015), their results are not used for comparisons in Table 2.11 (a comparison between RVND^{SQ}, RVND^{TO}, and all other solution methods is provided by Table 2.16 to Table 2.19 in Appendix C)

Table 2.11 Results of RVND^{SQ} and RVND^{TO} on instance sets CE and G

Instance	SM	BKOV	CT [s]	RVND^{SQ}			RVND^{TO}		
				OV	Imp. OV [%]	Imp. CT [%]	OV	Imp. OV [%]	Imp. CT [%]
CE-01	TS	1,119.47	24.3	1,135.81	-1.46	87.1	1,126.24	-0.60	94.0
CE-02	TS	1,814.52	33.0	1,841.69	-1.50	79.9	1,835.45	-1.15	95.9
CE-03	AVNS	1,920.86	212.1	1,965.90	-2.34	88.9	1,967.35	-2.42	96.6
CE-04	AVNS	2,512.05	279.7	2,564.22	-2.08	79.5	2,565.11	-2.11	94.1
CE-05	AVNS	3,099.77	228.6	3,162.54	-2.03	60.8	3,171.29	-2.31	88.8
CE-06	TS+	1,207.47	25.2	1,209.62	-0.18	85.9	1,220.91	-1.11	94.4
CE-07	TS	2,006.52	32.7	2,049.29	-2.13	84.4	2,051.09	-2.22	95.7
CE-08	AVNS	2,052.05	253.1	2,087.58	-1.73	91.3	2,087.62	-1.73	97.2
CE-09	AVNS	2,432.51	259.0	2,451.90	-0.80	79.1	2,454.59	-0.91	94.1
CE-10	AVNS	3,391.35	201.0	3,445.89	-1.61	54.9	3,459.41	-2.01	87.2
CE-11	AVNS	2,332.21	316.0	2,338.83	-0.28	83.1	2,339.11	-0.30	94.9
CE-12	TS+	1,952.86	59.5	1,970.87	-0.92	71.8	1,956.71	-0.20	92.1
CE-13	AVNS	2,858.94	278.5	2,864.16	-0.18	81.0	2,868.93	-0.35	94.1
CE-14	AVNS	2,215.38	93.2	2,224.25	-0.40	81.9	2,234.57	-0.87	94.7
Mean CE					-1.26	79.3		-1.31	93.8
G-01	AVNS	14,157.08	652.6	14,283.09	-0.89	-5.6	14,295.11	-0.97	47.9
G-02	AVNS	19,204.36	1558.4	19,394.37	-0.99	-52.2	19,436.23	-1.21	25.1
G-03	TS+	24,592.18	5940.4	24,824.57	-0.94	-28.9	24,834.50	-0.99	36.8
G-04	AVNS	34,415.82	2500.9	34,793.08	-1.10	-520.7	34,841.65	-1.24	-168.5
G-05	TS+	14,261.31	847.8	14,489.59	-1.60	-57.2	14,572.95	-2.19	36.5
G-06	AVNS	21,440.79	1783.5	21,684.74	-1.14	-79.2	21,820.27	-1.77	30.5
G-07	AVNS	23,375.60	2262.8	23,767.21	-1.68	-168.1	23,739.21	-1.56	-2.6
G-08	AVNS	29,797.62	2339.7	30,079.61	-0.95	-301.2	30,082.38	-0.96	-58.2
G-09	TS+	1,325.62	819.1	1,358.61	-2.49	33.0	1,354.38	-2.17	76.3
G-10	TS+	1,590.82	1762.3	1,621.51	-1.93	34.5	1,633.94	-2.71	76.5
G-11	TS	2,172.28	1492.7	2,218.96	-2.15	-67.6	2,222.61	-2.32	38.5
G-12	TS	2,492.75	2309.7	2,549.76	-2.29	-117.0	2,540.85	-1.93	21.0
G-13	TS+	2,274.12	504.5	2,326.12	-2.29	78.9	2,326.26	-2.29	92.9
G-14	TS+	2,703.31	976.9	2,792.45	-3.30	75.5	2,802.60	-3.67	91.6
G-15	TS	3,158.92	924.8	3,270.84	-3.54	49.4	3,280.20	-3.84	82.8

G-16	TS+	3,638.39	4675.1	3,779.11	-3.87	81.4	3,772.98	-3.70	93.6
G-17	TS+	1,633.35	472.0	1,658.32	-1.53	76.4	1,657.05	-1.45	92.2
G-18	TS	2,705.90	622.0	2,708.61	-0.10	70.7	2,729.05	-0.86	90.6
G-19	TS	3,497.54	1012.5	**3,473.82**	0.68	73.2	3,501.38	-0.11	91.1
G-20	TS+	4,306.89	2476.3	4,321.60	-0.34	85.2	4,335.79	-0.67	95.1
Mean G					-1.62	-37.0		-1.83	44.5
Mean					-1.47	10.9		-1.61	64.8

Table 2.12 Results of RVND^{SQ} and RVND^{TO} on the instance sets CE-H and G-H

Instance	SM	BKOV	CT [s]	RVND^{SQ}			RVND^{TO}		
				OV	Imp. OV [%]	Imp. CT [%]	OV	Imp. OV [%]	Imp. CT [%]
CE-H-01	TS	1,191.70	25.7	1,210.28	-1.56	86.7	1,199.03	-0.61	93.8
CE-H-02	TS+	1,791.21	34.8	1,828.29	-2.07	85.1	1,834.25	-2.40	95.6
CE-H-03	TS+	1,917.96	80.8	1,944.63	-1.39	70.8	1,936.65	-0.97	91.1
CE-H-04	TS	2,481.64	195.6	2,539.38	-2.33	68.5	2,500.89	-0.78	92.3
CE-H-05	TS+	3,143.01	342.4	3,219.14	-2.42	76.9	3,195.91	-1.68	94.6
CE-H-06	TS+	1,206.82	25.1	1,218.85	-1.00	87.0	1,220.39	-1.12	95.3
CE-H-07	TS+	2,031.85	32.0	2,053.31	-1.06	85.2	2,061.65	-1.47	96.4
CE-H-08	TS+	1,986.51	84.5	2,024.93	-1.93	66.7	2,024.96	-1.94	92.0
CE-H-09	TS	2,445.49	188.9	2,502.81	-2.34	71.2	2,497.76	-2.14	93.2
CE-H-10	TS	3,271.70	309.5	3,304.86	-1.01	66.9	3,331.62	-1.83	92.3
CE-H-11	RIP	2,308.76	18.8	2,323.08	-0.62	-183.3	2,331.63	-0.99	30.5
CE-H-12	RIP	1,908.74	13.0	1,920.32	-0.61	-30.8	1,918.97	-0.54	70.1
CE-H-13	RIP	2,842.18	19.5	2,855.39	-0.46	-156.6	2,882.67	-1.42	33.7
CE-H-14	TS+	1,907.75	67.2	1,914.08	-0.33	71.1	1,917.14	-0.49	93.1
Mean CE-H					-1.37	33.3		-1.31	83.1
G-H-01	TS	14,174.27	642.5	14,191.65	-0.12	-17.5	14,197.28	-0.16	55.6
G-H-02	TS+	18,537.70	4,955.3	18,541.59	-0.02	41.2	18,571.03	-0.18	78.0
G-H-03	TS+	25,177.92	11,996.2	25,297.58	-0.48	39.4	25,306.09	-0.51	75.6
G-H-04	TS+	34,991.21	4,079.1	**34,794.70**	0.56	-284.4	**34,747.39**	0.70	-55.7
G-H-05	TS+	15,411.82	754.2	15,714.24	-1.96	-50.1	15,751.18	-2.20	41.1
G-H-06	TS+	19,859.30	1,954.1	20,116.24	-1.29	-53.4	20,204.73	-1.74	41.2
G-H-07	TS+	23,481.28	9,167.3	23,751.16	-1.15	33.3	23,716.72	-1.00	72.4
G-H-08	TS+	27,334.84	18,625.2	27,482.48	-0.54	36.0	27,433.18	-0.36	75.1
G-H-09	TS+	1,329.27	1,829.3	1,373.90	-3.36	74.2	1,371.14	-3.15	90.9
G-H-10	TS	1,554.96	1,087.1	1,611.09	-3.61	-13.4	1,604.15	-3.16	55.0
G-H-11	TS	2,191.23	1,445.5	2,246.13	-2.51	-38.4	2,244.60	-2.44	45.9

G-H-12	TS+	2,482.92	4,224.0	2,554.43	-2.88	-7.2	2,559.52	-3.09	60.0
G-H-13	TS	2,231.88	405.8	2,307.28	-3.38	75.0	2,300.13	-3.06	91.4
G-H-14	TS+	2,684.70	1,042.9	2,752.90	-2.54	78.8	2,767.00	-3.07	92.8
G-H-15	TS	3,123.60	976.6	3,222.33	-3.16	52.6	3,237.13	-3.63	84.5
G-H-16	TS+	3,621.85	6,217.4	3,768.94	-4.06	87.7	3,770.16	-4.09	95.9
G-H-17	TS+	1,664.08	522.6	1,680.94	-1.01	80.8	1,685.81	-1.31	93.7
G-H-18	TS+	2,708.73	1,132.5	2,730.16	-0.79	86.5	2,727.52	-0.69	95.6
G-H-19	TS+	3,443.59	1,627.1	**3,439.96**	0.11	86.4	3,449.35	-0.17	95.3
G-H-20	TS	4,306.53	1,392.2	4,312.65	-0.14	80.2	4,334.54	-0.65	93.0
Mean G-H					-1.62	19.4		-1.70	68.9
Mean					-1.51	25.1		-1.54	74.7

The results of Table 2.11 and Table 2.12 show the general suitability of RVND^{SQ} and RVND^{TO} to solve VRPPCs. However, it has to be remarked that the effectiveness of RVND is inferior compared to the other solution methods considered here. Nevertheless, for the three instances G-19, G-H-04, and G-H-19, improvements of the best known objective values are achieved.

Regarding the trade-off between solution quality and computation time, it has to be emphasized that RVND^{TO} has a remarkable efficiency for almost every instance compared to the other solution methods: RVND^{TO} achieves for homogeneous and heterogeneous vehicle fleets small mean degradations of 1.61% and 1.54%, whereby computation times are reduced by 64.8% and 74.7%.

2.5.3 Economic benefits and implications of solving the VRPPCdR

Purpose of this part of the analysis is the assessment of the economic benefits (cost reductions) achieved by solving the VRPPCdR and the investigation of further managerial implications. Therefore, we compare the costs between the two planning approaches VRPPCdR and VRPPC with regard to the basic data scenarios provided by the BOL instance set. For the VRPPCdR instance sets N1 to N5, Table 2.13 first lists the objective value ("OV^{C}"), the relative gap reported by CPLEX ("Rel. Gap"), and the "used vehicles" for the solutions calculated by CPLEX 12.5 (terminated after 50 hours). Next columns contain the objective value ("$\text{OV}^{RVND\text{-}SQ}$"), the relative improvement of RVND^{SQ} comparing $\text{OV}^{RVND\text{-}SQ}$ and OV^{C} ("Rel. Imp. OV"), and the list of "used vehicles" for the best solution (out of ten) calculated by RVND^{SQ}. The column "Identical" indicates that the list of used vehicles and the costs are identical (√), the lists are identical and the costs achieved by RVND^{SQ} are lower (+), the lists are not identical and the costs achieved by RVND^{SQ} are lower (++), the lists are identical and the costs of RVND^{SQ} are higher (-), or the lists are not identical and the costs of RVND^{SQ} are higher (-). The last two columns compare the total cost ("TC") improvements of the RVND^{SQ} solutions of the VRPPCdR instance with the total costs ("$\text{TC}_{\text{BOL-1, .., 5}}$"; including fixed costs) of the original VRPPC instance (see first column of Table 2.14 in Appendix C).

The column "used vehicles" shows that the private fleet is utilized before another delivery mode is used. This is in line with the assumption that FTL deliveries with the private fleet are more cost-efficient than using other delivery modes. Concerning this, we conclude that the parameters of the rental options are reasonable.

Comparing the solutions calculated by CPLEX and RVND^{SQ}, RVND^{SQ} selects almost for each problem instance the most suitable delivery modes (or even better ones) and makes only in two cases (N3-06 and N3-07) another mode selection decision than CPLEX leading to higher costs. Although the vehicle combination is identical, RVND^{SQ} fails in two cases (N3-04 and N4-07) to allocate the customers and route the vehicles most preferably. Based on these observations, it can be stated that the RVND^{SQ} solution method is able to determine delivery modes appropriately, performs well in terms of solution quality, and requires an extraordinarily small computation time.

Table 2.13 Results of RVND^{SQ} on instance sets N1 to N5

		CPLEX 12.5 (after 50 h)			RVND^{SQ}						
Basic data	Instance	OV^C	Rel. Gap [%]	Used vehicles	$OV^{RVND\text{-}SQ}$	Rel. Imp. OV [%]	CT [s]	Used vehicles	Identical	TC	Rel. Imp. TC [%]
N1 (TC_{BOL-1}=423.5)	-01	333.76	0.0	V1;C2	333.76	0.0	0.2	V1;C2	√	393.76	7.0
	-02	332.64	0.0	V1;C1;C2	332.64	0.0	0.4	V1;C1;C2	√	392.64	7.3
	-03	234.00	0.0	V1;D2	234.00	0.0	0.4	V1;D2	√	294.00	30.6
	-04	234.00	0.0	V1;D2	234.00	0.0	0.6	V1;D2	√	294.00	30.6
	-05	234.00	0.0	V1;D2	234.00	0.0	0.7	V1;D2	√	294.00	30.6
	-06	234.00	0.0	V1;D2	234.00	0.0	0.6	V1;D2	√	294.00	30.6
	-07	234.00	0.0	V1;D2	234.00	0.0	0.7	V1;D2	√	294.00	30.6
N2 (TC_{BOL-2}=476.5)	-01	431.52	0.0	V2;C2;C3	431.52	0.0	0.2	V2;C2;C3	√	551.52	-15.7
	-02	385.52	0.0	V2;C4	385.52	0.0	0.6	V2;C4	√	505.52	-6.1
	-03	298.50	0.0	V2;M1	298.50	0.0	1.3	V2;M1	√	418.50	12.2
	-04	293.00	0.0	V2;M2;C1	293.00	0.0	1.3	V2;M2;C1	√	413.00	13.3
	-05	304.54	0.0	V2;M2;C4	304.54	0.0	1.0	V2;M2;C4	√	424.54	10.9
	-06	372.97	0.0	V2;D2;C3	372.97	0.0	1.3	V2;D2;C3	√	492.97	-3.5
	-07	293.00	0.0	V2;M2;C1	293.00	0.0	2.0	V2;M2;C1	√	413.00	13.3
N3 (TC_{BOL-3}=777.0)	-01	501.49	25.3	V3;V4;C3	501.49	0.0	0.8	V3;V4;C3	√	791.49	-1.9
	-02	456.00	20.4	V3;V4;C1	456.00	0.0	1.5	V3;V4;C1	√	746.00	4.0
	-03	393.00	37.4	V3;V4;M1	385.00	2.0	1.7	V3;V4;M1	+	675.00	13.1
	-04	360.00	0.0	V3;V4;M2	395.25	-9.8	1.7	V3;V4;M2	-	685.25	11.8
	-05	360.00	0.0	V3;V4;M2	360.00	0.0	1.3	V3;V4;M2	√	650.00	16.3
	-06	490.50	57.8	V3;V4;D2	501.49	-2.2	2.2	V3;V4;C3	--	791.49	-1.9
	-07	360.00	19.4	V3;V4;M2	385.50	-7.1	0.7	V3;V4;M1	--	675.50	13.1
N4 (TC_{BOL-4}=1515.0)	-01	1802.75	48.1	V5;C2;C3	1339.24	25.7	1.0	V5;C3	++	1589.24	-4.9
	-02	1278.00	26.3	V5;C1;C3	1278.00	0.0	0.3	V5;C1;C3	√	1528.00	-0.9
	-03	1056.00	43.9	V5;M1	1043.50	1.2	2.3	V5;M1;M3	++	1293.50	14.6
	-04	1112.50	26.3	V5;M3;C1	1082.50	2.7	1.9	V5;M3;C1	+	1332.50	12.0
	-05	1385.75	35.7	V5;M2;D2	1185.14	14.5	1.5	V5;M2;C3	++	1435.14	5.3
	-06	1400.00	59.3	V5;M3;D2	1141.14	18.5	1.6	V5;M3;C3	++	1391.14	8.2
	-07	1043.50	44.4	V5;M1;M3	1059.50	-1.5	2.6	V5;M1;M3	-	1309.50	13.6

N5 (TC_{BOL-5}=1617.0)	-01	1494.48	50.6	V5;V6;C2;C3	1133.05	24.2	0.8	V5;V6;C3	++	1333.05	17.6
	-02	1073.50	29.9	V5;V6;C1	1073.50	0.0	1.5	V5;V6;C1	√	1273.50	21.2
	-03	938.50	61.8	V5;V6;M1	938.50	0.0	1.7	V5;V6;M1	√	1138.50	29.6
	-04	1033.00	46.4	V5;V6;M3;C1	940.50	9.0	2.5	V5;V6;M2	++	1140.50	29.5
	-05	958.50	41.0	V5;V6;M2	943.50	1.6	2.0	V5;V6;M2	+	1143.50	29.3
	-06	1389.42	73.2	V5;V6;M3;C2	1028.14	26.0	2.8	V5;V6;M3;C3	++	1228.14	24.0
	-07	1091.00	65.6	V5;V6;M3;C1	938.50	14.0	1.4	V5;V6;M1	++	1138.50	29.6
Mean						3.7	1.3				13.6

Regarding the economic potentials of the VRPPCdR approach, the mean relative cost improvement of 13.6% is impressive. Nevertheless, for some instances no cost reduction can be achieved (e. g., N2-01). For these cases, the downsizing of the private vehicle fleets (see section 2.5.1) cannot be compensated by the newly available outsourcing options. This fact and the comparison of the used vehicles lead to the conclusion that the available outsourcing possibilities (and their specific parameters) are directly related to the cost reduction potential. In this context, it has to be highlighted that the largest cost reduction for a specific basic data scenario (BOL-1 to BOL-5) is not only achieved with the carrier and rental option scenario S7 (including all possible rental options) but also with carrier and rental option scenarios offering only a subset of all possible rental options (e. g., S4 for BOL-2 or S5 for BOL-4). Because different carrier and rental option scenarios (all with the same parameters) achieve this for different basic data scenarios, we conclude that there is not only one most preferable set of outsourcing options but it depends on basic data scenario.

Summarizing this analysis, it is shown that the VRPPCdR approach leads to notable cost reductions, that the best combination of private fleet and subcontracting options depends on the basic data scenario, and that the available subcontracting possibilities should be considered during fleet size planning to establish a reasonable equipped private fleet.

2.6 Conclusions and further research

In this paper, we tackle a very comprehensive delivery planning problem for shipping companies. It integrates two types of rental options, one based on mileage costs and one based on daily costs, a carrier-dependent cost function considering volumes and distances, and common carriers that offer volume discounts. All these aspects are considered simultaneously for the first time and define the new VRPPCdR. The presented MIP specifies the problem in detail and can be used by standard solvers to solve small instances to optimality. Because of the NP-hardness of the VRPPCdR and the goal to solve problem instances of virtually any size, several solution methods, based on the principles of VNS are developed: VNS, RVND, and VNSac. The evaluation of these methods shows that RVND using the introduced explicit shaking mechanism (including the RD-transformation rule) is superior to the other proposed methods. Furthermore, the invented distance proportionate vehicle selection mechanism is able to further increase the efficiency of RVND. In addition, the decision-maker is able to balance between solution quality and computation time by adjusting a single parameter.

The economic potentials of the VRPPCdR approach are assessed based on 35 problem instances and a remarkable mean cost reduction by 13.6% is reported. Furthermore, we have shown a dependency between the available subcontracting possibilities, basic data scenario, and the economic potentials for cost reduction. In addition, the influence between available subcontracting possibilities on the tactical fleet size planning problem is carved out.

This dependency between the available subcontracting possibilities on the fleet size planning problem suggests further research opportunities. Also, an improvement of the proposed solution methods for very large problem instances, particularly instances with long routes, appears to be a promising issue. For example, ejection chain transformations or the adaptive mechanism proposed by Stenger et al. (2013b) could be integrated. Moreover, the integration of one of the proposed VNS-variants as local search procedure within a population based metaheuristic like a Genetic Algorithm or Ant Colony Optimization seem to be worth to be investigated in future. For example, Nagata and Bräysy (2009) and Böhnlein et al. (2009) have shown the suitability of such approaches to solve VRPs or VRP variants. In this context, problem specific operators combining solutions with different combinations of used vehicles and common carriers might be interesting.

In addition, the idea of the vehicle selection mechanism for inter-route optimization procedures based on route characteristics (like the absolute distance of their center points) could be extended in the future: For example by considering their relative or absolute position to each other. Also the relative or absolute position of customers could be used to increase the probability of improving λ-interchanges.

Beside the investigation of other solution methods or their components, the problem definition itself could be enhanced by adding characteristics such as time windows or multiple depots.

Appendix A: Linearization of the last part of objective function (5)

In a first step, we introduce an auxiliary decision variable B_l that indicates whether the minimal transportation rate tr_l^{Min} is used.

$$B_l = \begin{cases} 1 \quad \text{if } tr_l - \left(df_l \cdot \sum_{j \in C^n} Z_{jl} \cdot q_j \right) \geq \mathrm{tr}_l^{Min} \\ 0 \qquad\qquad \text{otherwise} \end{cases}$$

This new variable leads to the following transformation of the last summand of (5):

$$\left(\sum_{l \in E} \sum_{i \in C^n} Z_{il} \cdot B_l \cdot \left(tr_l - \left(df_l \cdot \sum_{j \in C^n} Z_{jl} \cdot q_j \right) \right) \cdot \omega_{il} \right) + \left(\sum_{l \in E} \sum_{i \in C^n} Z_{il} \cdot (1 - B_l) \cdot tr_l^{Min} \cdot \omega_{il} \right) \tag{27}$$

$$\begin{aligned} &= \left(\sum_{l \in E} \sum_{i \in C^n} Z_{il} \cdot B_l \cdot tr_l \cdot \omega_{il} \right) - \left(\sum_{l \in E} \sum_{i \in C^n} \sum_{j \in C^n} Z_{il} \cdot B_l \cdot df_l \cdot \omega_{il} \cdot Z_{jl} \cdot q_j \right) \\ &\qquad + \left(\sum_{l \in E} tr_l^{Min} \sum_{i \in C^n} \left(Z_{il} - B_l \cdot Z_{il} \right) \cdot \omega_{il} \right) \end{aligned} \tag{28}$$

However, introducing B_l eliminates the maximum operator, but (27) and also (28) remain non-linear. To linearize (28), the product of the decision variables (Z_{il}, Z_{jl}, and B_l) is substituted by two additional auxiliary variables: P_{il} substitutes the product of B_l and Z_{il}; Q_{ijl} the product of B_l, Z_{il}, and Z_{jl}.

$$P_{il} = \begin{cases} 1 \quad \text{if } Z_{il} = 1 \text{ and } \mathrm{B}_l = 1 \\ 0 \quad \text{otherwise} \end{cases}$$

$$Q_{ijl} = \begin{cases} 1 \quad \text{if } Z_{jl} = P_{il} = 1, \text{which implicates } Z_{il} = B_l = 1 \\ 0 \quad \text{otherwise} \end{cases}$$

After incorporating P_{il} and Q_{ijl}, formula (28) can be written as follows (note that B_l is eliminated):

$$\left(\sum_{l \in E} tr_l \sum_{i \in C^n} P_{il} \cdot \omega_{il} \right) - \left(\sum_{l \in E} df_l \sum_{i \in C^n} \sum_{j \in C^n} Q_{ijl} \cdot \omega_{il} \cdot q_j \right) + \left(\sum_{l \in E} tr_l^{Min} \sum_{i \in C^n} \left(Z_{il} - P_{il} \right) \cdot \omega_{il} \right) \tag{29}$$

Appendix B: Illustration of the common carriers' cost functions

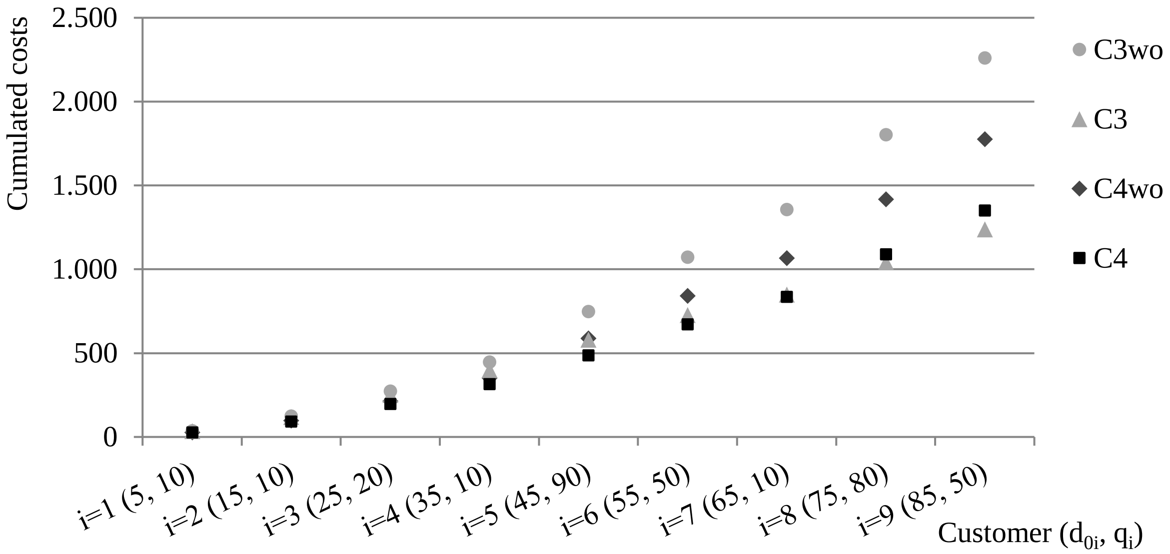


Fig. 2.2 Cumulated transportation costs with volume discounts (curves C3 and C4) and without volume discounts (curves C3wo and C4wo).

Appendix C: Tables

Table 2.14 Reference objective values and relative gaps for instance sets CHU and BOL

Instance	Objective value without fixed costs	Relative gap of CPLEX after 50 hours [%]	Fixed costs	Total costs TC
CHU-1	277.5	0.0	110	387.5
CHU-2	366.0	0.0	220	586.0
CHU-3	403.5	0.0	420	823.5
CHU-4	939.0	6.5	450	1389.0
CHU-5	835.5	37.3	630	1465.5
BOL-1	313.5	0	110	423.5
BOL-2	256.5	0	220	476.5
BOL-3	357.0	0	420	777.0
BOL-4	1065.0	13.1	450	1515.0
BOL-5	987.0	27.6	630	1617.0

Table 2.15: Basic data with location (x, y), demand quantities q_i, and distance to depot d_{i0}

Bolduc 1

i	x	y	q_i	d_{i0}
0	35	35		
1	28	59	10	25
2	35	42	7	7
3	45	53	13	20
4	54	22	19	23
5	68	63	26	43

Bolduc 4

i	x	y	q_i	d_{i0}
0	266	235		
1	227	276	125	56
2	303	243	84	37
3	312	196	60	60
4	258	196	500	39
5	286	195	300	44
6	204	186	175	79
7	249	212	350	28
8	209	268	150	65
9	323	212	1100	61
10	299	267	4100	45
11	300	254	225	38
12	312	225	300	47
13	300	268	250	47
14	305	278	500	58
15	251	238	150	15
16	294	238	100	28
17	256	192	250	44
18	293	202	120	42
19	291	202	600	41
20	221	186	500	66
21	317	181	175	74
22	265	206	75	29

Bolduc 2

i	x	y	q_i	d_{i0}
0	30	40		
1	40	33	7	12
2	26	43	30	5
3	47	60	16	26
4	46	26	9	21
5	67	25	21	39
6	52	57	15	27
7	38	31	19	12
8	36	59	23	19
9	50	39	11	20
10	47	22	5	24

Bolduc 3

i	x	y	q_i	d_{i0}
0	40	40		
1	44	43	18	5
2	42	59	26	19
3	47	57	11	18
4	29	55	30	18
5	29	30	21	14
6	46	39	19	6
7	67	22	15	32
8	64	65	16	34
9	45	52	29	13
10	65	68	26	37
11	33	29	37	13
12	64	24	16	28
13	55	66	12	30
14	44	54	31	14
15	41	22	8	18

Bolduc 5

i	x	y	q_i	d_{i0}
0	162	354		
1	111	354	300	51
2	136	355	3100	26
3	183	401	125	51
4	214	374	100	55
5	131	371	200	35
6	215	332	150	57
7	141	385	150	37
8	173	334	450	22
9	156	388	300	34
10	164	352	100	2
11	126	341	950	38
12	104	402	125	75
13	171	346	150	12
14	146	393	150	42
15	201	393	550	55
16	118	354	150	44
17	149	337	100	21
18	221	397	150	73
19	117	385	400	54
20	138	403	300	54
21	194	409	1500	63
22	177	361	100	16
23	163	336	300	18
24	158	387	500	33
25	180	392	800	42
26	119	399	300	62
27	128	352	100	34
28	116	346	150	46
29	140	407	1000	57

Table 2.16: CE-instances: Minimum objective values (“OV”), mean computation times (“CT”), and relative improvements (“Gap”; OV vs. BKOV)

	RIP (Intel Xeon, 3.6 Ghz)			TS (AMD Opteron 275, 2.2 Ghz)			TS+ (AMD Opteron 275, 2.2 Ghz)			AVNS (Intel Core i5, 2.67 GHz)			UHGS (Intel Xeon, 3.07 GHz, mean CT: 1,583.4 [s])		MS-ILS (Intel Xeon, 3.07 GHz, mean CT: 996.6 [s])		MS-LS (Intel Xeon, 3.07 GHz, mean CT: 113.4 [s])			RVND^{SQ} (Intel Xeon, 2.9 GHz)			RVND^{TO} (Intel Xeon, 2.9 GHz)		
Inst.	OV	Gap [%]	CT [s]	OV	Gap [%]	CT [s]	OV	Gap [%]	CT [s]	OV	Gap [%]	CT [s]	OV	Gap [%]	OV	Gap [%]	OV	Gap [%]	BKOV	OV	Gap [%]	CT [s]	OV	Gap [%]	CT [s]
CE-01	1,132.91	-1.20	25.0	**1,119.47**	0.00	24.3	**1,119.47**	0.00	24.9	1,123.95	-0.40	92.5	**1,119.47**	0.00	**1,119.47**	0.00	1,128.28	-0.79	1,119.47	1,135.81	-1.46	3.1	1,126.24	-0.60	1.5
CE-02	1,835.76	-1.17	73.0	**1,814.52**	0.00	33.0	**1,814.52**	0.00	33.9	**1,814.52**	0.00	48.6	**1,814.52**	0.00	**1,814.52**	0.00	1,883.44	-3.80	1,814.52	1,841.69	-1.50	6.6	1,835.45	-1.15	1.4
CE-03	1,959.65	-2.12	107.0	1,921.10	-0.11	78.6	1,930.66	-0.60	81.0	1,920.86	-0.09	212.1	**1,919.05**	0.00	1,922.18	-0.16	1,958.80	-2.07	1,919.05	1,965.90	-2.44	23.5	1,967.35	-2.52	7.2
CE-04	2,545.72	-1.61	250.0	2,525.17	-0.79	193.2	2,525.17	-0.79	200.6	2,512.05	-0.27	279.7	**2,505.39**	0.00	**2,505.39**	0.00	2,568.49	-2.52	2,505.39	2,564.22	-2.35	57.3	2,565.11	-2.38	16.4
CE-05	3,172.22	-2.94	474.0	3,113.58	-1.04	309.9	3,117.10	-1.15	353.2	3,099.77	-0.59	228.6	**3,081.59**	0.00	3,090.53	-0.29	3,201.29	-3.88	3,081.59	3,162.54	-2.63	89.5	3,171.29	-2.91	25.6
CE-06	1,208.33	-0.07	25.0	**1,207.47**	0.00	25.5	**1,207.47**	0.00	25.2	1,207.81	-0.03	75.9	**1,207.47**	0.00	**1,207.47**	0.00	1,216.57	-0.75	1,207.47	1,209.62	-0.18	3.5	1,220.91	-1.11	1.4
CE-07	**2,006.52**	0.00	71.0	**2,006.52**	0.00	32.7	**2,006.52**	0.00	34.0	2,013.93	-0.37	50.9	**2,006.52**	0.00	**2,006.52**	0.00	2,079.67	-3.65	2,006.52	2,049.29	-2.13	5.1	2,051.09	-2.22	1.4
CE-08	2,082.75	-1.50	110.0	2,060.17	-0.40	85.1	2,056.59	-0.22	81.6	**2,052.05**	0.00	253.1	**2,052.05**	0.00	2,054.64	-0.13	2,100.59	-2.37	2,052.05	2,087.58	-1.73	21.9	2,087.62	-1.73	7.1
CE-09	2,443.94	-0.77	260.0	2,438.43	-0.54	185.3	2,435.97	-0.44	188.2	2,432.51	-0.30	259.0	**2,425.32**	0.00	2,428.03	-0.11	2,505.24	-3.30	2,425.32	2,451.90	-1.10	54.0	2,454.59	-1.21	15.3
CE-10	3,464.90	-2.46	478.0	3,406.82	-0.74	311.1	3,401.83	-0.60	345.7	3,391.35	-0.29	201.0	**3,381.67**	0.00	3,382.23	-0.02	3,491.59	-3.25	3,381.67	3,445.89	-1.90	90.6	3,459.41	-2.30	25.8
CE-11	2,333.03	-0.09	195.0	2,353.39	-0.96	126.3	2,332.36	-0.06	131.0	2,332.21	-0.05	316.0	**2,330.94**	0.00	**2,330.94**	0.00	2,408.13	-3.31	2,330.94	2,338.83	-0.34	53.5	2,339.11	-0.35	16.2
CE-12	1,953.55	-0.04	128.0	**1,952.86**	0.00	60.4	**1,952.86**	0.00	59.5	1,953.55	-0.04	92.9	**1,952.86**	0.00	**1,952.86**	0.00	1,982.06	-1.50	1,952.86	1,970.87	-0.92	16.8	1,956.71	-0.20	4.7
CE-13	2,864.21	-0.19	188.0	2,882.70	-0.83	130.0	2,860.89	-0.07	132.1	2,858.94	0.00	278.5	**2,858.83**	0.00	**2,858.83**	0.00	3,025.26	-5.82	2,858.83	2,864.16	-0.19	52.8	2,868.93	-0.35	16.5
CE-14	2,224.63	-0.52	110.0	2,216.97	-0.18	65.0	2,216.97	-0.18	64.2	2,215.38	-0.11	93.2	**2,213.02**	0.00	**2,213.02**	0.00	2,226.44	-0.61	2,213.02	2,224.25	-0.51	16.9	2,234.57	-0.97	5.0
Mean		-1.05	178.1		-0.40	118.6		-0.29	125.4		-0.18	177.3		0.00		-0.05		-2.69			-1.38	35.4		-1.43	10.4

Table 2.17: G-instances: Minimum objective values ("OV"), mean computation times ("CT"), and relative improvements ("Gap"; OV vs. BKOV)

	RIP (Intel Xeon, 3.6 Ghz)			TS (AMD Opteron 275, 2.2 Ghz)			TS+ (AMD Opteron 275, 2.2 Ghz)			AVNS (Intel Core i5, 2.67 GHz)			UHGS (Intel Xeon, 3.07 GHz, mean CT: 1,583.4 [s])		MS-ILS (Intel Xeon, 3.07 GHz, mean CT: 996.6 [s])		MS-LS (Intel Xeon, 3.07 GHz, mean CT: 113.4 [s])			RVND^{SQ} (Intel Xeon, 2.9 GHz)			RVND^{TO} (Intel Xeon, 2.9 GHz)		
Inst.	OV	Gap [%]	CT [s]	OV	Gap [%]	CT [s]	OV	Gap [%]	CT [s]	OV	Gap [%]	CT [s]	OV	Gap [%]	OV	Gap [%]	OV	Gap [%]	BKOV	OV	Gap [%]	CT [s]	OV	Gap [%]	CT [s]
G-01	14,388.58	-1.68	651.0	14,218.83	-0.48	638.3	14,190.01	-0.27	1,183.8	14,157.08	-0.04	652.6	**14,151.51**	0.00	14,165.45	-0.10	14,329.96	-1.26	14,151.51	14,283.09	-0.93	689.1	14,295.11	-1.01	340.1
G-02	19,505.00	-1.64	1,178.0	19,729.96	-2.81	1,215.2	19,208.52	-0.09	5,220.6	19,204.36	-0.07	1,558.4	**19,190.77**	0.00	19,191.56	0.00	19,524.50	-1.74	19,190.77	19,394.37	-1.06	2,371.7	19,436.23	-1.28	1,166.9
G-03	24,978.17	-1.59	2,061.0	25,653.58	-4.33	2,241.3	24,592.18	-0.02	5,940.4	24,602.61	-0.06	2,356.1	**24,588.29**	0.00	24,609.36	-0.09	25,038.41	-1.83	24,588.29	24,824.57	-0.96	7,654.6	24,834.50	-1.00	3,751.4
G-04	34,957.98	-1.58	3,027.0	36,022.73	-4.67	3,833.9	34,802.08	-1.12	5,508.7	**34,415.82**	0.00	2,500.9	34,517.47	-0.30	34,907.49	-1.43	35,182.78	-2.23	34,415.82	34,793.08	-1.10	15,522.5	34,841.65	-1.24	6,715.4
G-05	14,683.03	-2.96	589.0	14,673.56	-2.89	875.0	**14,261.31**	0.00	847.8	14,272.32	-0.08	1,301.1	14,296.07	-0.24	14,373.87	-0.79	14,735.12	-3.32	14,261.31	14,489.59	-1.60	1,332.4	14,572.95	-2.19	538.7
G-06	22,260.19	-3.82	1,021.0	22,278.99	-3.91	1,445.3	21,498.03	-0.27	1,591.1	**21,440.79**	0.00	1,783.5	21,488.29	-0.22	21,546.18	-0.49	22,024.07	-2.72	21,440.79	21,684.74	-1.14	3,196.3	21,820.27	-1.77	1,239.5
G-07	23,963.36	-2.51	1,628.0	24,191.41	-3.49	2,052.8	23,513.06	-0.59	5,514.0	**23,375.60**	0.00	2,262.8	23,463.05	-0.37	23,547.12	-0.73	23,980.00	-2.59	23,375.60	23,767.21	-1.68	6,065.5	23,739.21	-1.56	2,321.7
G-08	30,496.18	-2.34	2,419.0	30,627.91	-2.79	3,059.9	30,073.56	-0.93	5,729.0	**29,797.62**	0.00	2,339.7	29,918.06	-0.40	30,064.28	-0.89	30,459.11	-2.22	29,797.62	30,079.61	-0.95	9,385.7	30,082.38	-0.96	3,700.4
G-09	1,341.17	-1.17	832.0	1,328.14	-0.19	611.0	**1,325.62**	0.00	819.1	1,335.45	-0.74	602.0	1,332.63	-0.53	1,339.06	-1.01	1,397.08	-5.39	1,325.62	1,358.61	-2.49	549.1	1,354.38	-2.17	193.7
G-10	1,612.09	-1.34	1,294.0	1,590.83	0.00	938.8	**1,590.82**	0.00	1,762.3	1,604.50	-0.86	978.4	1,603.82	-0.82	1,617.58	-1.68	1,682.31	-5.75	1,590.82	1,621.51	-1.93	1,154.7	1,633.94	-2.71	413.9
G-11	2,198.45	-1.20	2,004.0	**2,172.28**	0.00	1,492.7	2,173.80	-0.07	3,284.3	2,189.02	-0.77	1,534.3	2,192.68	-0.94	2,228.23	-2.58	2,281.79	-5.04	2,172.28	2,218.96	-2.15	2,502.4	2,222.61	-2.32	918.6
G-12	2,521.79	-1.16	2,900.0	**2,492.75**	0.00	2,309.7	2,495.02	-0.09	8,587.6	2,520.29	-1.10	2,043.9	2,529.84	-1.49	2,553.40	-2.43	2,652.57	-6.41	2,492.75	2,549.76	-2.29	5,010.9	2,540.85	-1.93	1,824.7
G-13	2,286.91	-1.12	802.0	2,278.99	-0.77	360.8	2,274.12	-0.56	504.5	2,291.83	-1.34	116.5	**2,261.50**	0.00	2,277.57	-0.71	2,337.43	-3.36	2,261.50	2,326.12	-2.86	106.4	2,326.26	-2.86	36.0
G-14	2,750.75	-2.35	1,251.0	2,705.00	-0.65	610.4	2,703.31	-0.59	976.9	2,708.22	-0.77	183.5	**2,687.50**	0.00	2,708.56	-0.78	2,791.23	-3.86	2,687.50	2,792.45	-3.91	239.3	2,802.60	-4.28	82.1
G-15	3,216.99	-2.06	1,862.0	3,158.92	-0.22	924.8	3,161.26	-0.29	1,952.0	3,194.82	-1.36	357.3	**3,152.00**	0.00	3,177.53	-0.81	3,296.86	-4.60	3,152.00	3,270.84	-3.77	467.9	3,280.20	-4.07	159.4
G-16	3,693.62	-1.70	2,778.0	3,639.11	-0.19	1,313.7	3,638.39	-0.17	4,675.1	3,671.34	-1.08	561.2	**3,632.04**	0.00	3,672.62	-1.12	3,811.80	-4.95	3,632.04	3,779.11	-4.05	869.5	3,772.98	-3.88	297.6
G-17	1,701.58	-4.18	806.0	1,636.11	-0.17	402.6	**1,633.35**	0.00	472.0	1,682.49	-3.01	110.3	1,671.72	-2.35	1,677.37	-2.70	1,717.55	-5.16	1,633.35	1,658.32	-1.53	111.4	1,657.05	-1.45	36.7
G-18	2,765.92	-2.22	1,303.0	**2,705.90**	0.00	622.0	2,710.21	-0.16	892.0	2,741.80	-1.33	156.4	2,733.12	-1.01	2,741.10	-1.30	2,801.70	-3.54	2,705.90	2,708.61	-0.10	182.5	2,729.05	-0.86	58.8
G-19	3,576.92	-2.27	1,903.0	**3,497.54**	0.00	1,012.5	3,497.72	-0.01	1,418.4	3,507.94	-0.30	194.1	3,504.26	-0.19	3,515.47	-0.51	3,585.64	-2.52	3,497.54	**3,473.82**	0.68	271.7	3,501.38	-0.11	89.8
G-20	4,378.13	-1.65	2,800.0	4,311.17	-0.10	1,368.9	**4,306.89**	0.00	2,476.3	4,332.44	-0.59	290.3	4,319.37	-0.29	4,333.59	-0.62	4,433.41	-2.94	4,306.89	4,321.60	-0.34	366.3	4,335.79	-0.67	120.9
Mean		-2.03	1,655.5		-1.38	1,366.5		-0.26	2,967.8		-0.68	1,094.2		-0.46		-1.04		-3.57			-1.71	2,902.5		-1.92	1,200.3

Table 2.18: CE-H-instances: Minimum objective values ("OV"), mean computation times ("CT"), and relative improvements ("Gap"; OV vs. BKOV)

	RIP (Intel Xeon 3.6, Ghz)			TS (AMD Opteron 275, 2.2 Ghz)			TS+ (AMD Opteron 275, 2.2 Ghz)				RVND^{SQ} (Intel Xeon, 2.9 GHz)			RVND^{TO} (Intel Xeon, 2.9 GHz)		
Inst.	OV	Gap [%]	CT [s]	OV	Gap [%]	CT [s]	OV	Gap [%]	CT [s]	BKOV	OV	Gap [%]	CT [s]	OV	Gap [%]	CT [s]
CE-H-01	1,192.72	-0.09	2.6	**1,191.70**	0.00	25.7	**1,191.70**	0.00	26.0	1,191.70	1,210.28	-1.56	3.4	1,199.03	-0.61	1.6
CE-H-02	1,798.26	-0.39	7.2	1,795.51	-0.24	33.7	**1,791.21**	0.00	34.8	1,791.21	1,828.29	-2.07	5.2	1,834.25	-2.40	1.5
CE-H-03	1,934.85	-0.88	10.5	1,926.33	-0.44	79.0	**1,917.96**	0.00	80.8	1,917.96	1,944.63	-1.39	23.6	1,936.65	-0.97	7.2
CE-H-04	2,493.93	-0.50	25.1	**2,481.64**	0.00	195.6	2,481.68	0.00	198.5	2,481.64	2,539.38	-2.33	61.6	2,500.89	-0.78	15.2
CE-H-05	3,195.66	-1.68	49.0	3,143.92	-0.03	295.9	**3,143.01**	0.00	342.4	3,143.01	3,219.14	-2.42	79.1	3,195.91	-1.68	18.6
CE-H-06	1,210.23	-0.28	2.5	**1,206.82**	0.00	25.4	**1,206.82**	0.00	25.1	1,206.82	1,218.85	-1.00	3.3	1,220.39	-1.12	1.2
CE-H-07	2,042.79	-0.54	7.4	2,035.90	-0.20	32.5	**2,031.85**	0.00	32.0	2,031.85	2,053.31	-1.06	4.7	2,061.65	-1.47	1.2
CE-H-08	2,015.72	-1.47	11.2	1,991.23	-0.24	81.4	**1,986.51**	0.00	84.5	1,986.51	2,024.93	-1.93	28.1	2,024.96	-1.94	6.8
CE-H-09	2,445.88	-0.02	26.7	**2,445.49**	0.00	188.9	2,447.58	-0.09	193.0	2,445.49	2,502.81	-2.34	54.4	2,497.76	-2.14	12.8
CE-H-10	3,304.69	-1.01	48.2	**3,271.70**	0.00	309.5	3,272.37	-0.02	342.4	3,271.70	3,304.86	-1.01	102.4	3,331.62	-1.83	24.0
CE-H-11	**2,308.76**	0.00	18.8	2,325.00	-0.70	127.0	2,336.51	-1.20	133.9	2,308.76	2,323.08	-0.62	53.3	2,331.63	-0.99	13.1
CE-H-12	**1,908.74**	0.00	13.0	1,912.47	-0.20	60.7	1,915.05	-0.33	60.4	1,908.74	1,920.32	-0.61	17.0	1,918.97	-0.54	3.9
CE-H-13	**2,842.18**	0.00	19.5	2,872.14	-1.05	125.0	2,868.13	-0.91	136.5	2,842.18	2,855.39	-0.46	50.0	2,882.67	-1.42	12.9
CE-H-14	1,920.36	-0.66	11.4	1,925.46	-0.93	65.8	**1,907.75**	0.00	67.2	1,907.75	1,914.08	-0.33	19.4	1,917.14	-0.49	4.7
Mean		-0.54	18.1		-0.29	117.6		-0.18	125.5			-1.37	36.1		-1.31	8.9

Table 2.19: G-H-instances: Minimum objective values ("OV"), mean computation times ("CT"), and relative improvements ("Gap"; OV vs. BKOV)

	RIP (Intel Xeon 3.6, Ghz)			TS (AMD Opteron 275, 2.2 Ghz)			TS+ (AMD Opteron 275, 2.2 Ghz)				RVND^{SQ} (Intel Xeon, 2.9 GHz)			RVND^{TO} (Intel Xeon, 2.9 GHz)		
Inst.	OV	Gap [%]	CT [s]	OV	Gap [%]	CT [s]	OV	Gap [%]	CT [s]	BKOV	OV	Gap [%]	CT [s]	OV	Gap [%]	CT [s]
G-H-01	14,408.31	-1.65	64.7	**14,174.27**	0.00	642.5	14,194.13	-0.14	1,193.7	14,174.27	14,191.65	-0.12	755.0	14,197.28	-0.16	285.5
G-H-02	18,663.15	-0.68	125.4	19,056.69	-2.80	1,284.9	**18,537.70**	0.00	4,955.3	18,537.70	18,541.59	-0.02	2,915.8	18,571.03	-0.18	1,092.4
G-H-03	25,561.55	-1.52	205.3	25,299.93	-0.48	2,258.7	**25,177.92**	0.00	11,996.2	25,177.92	25,297.58	-0.48	7,265.1	25,306.09	-0.51	2,925.7
G-H-04	35,495.66	-1.44	290.4	35,388.06	-1.13	4,095.7	**34,991.21**	0.00	4,079.1	34,991.21	**34,794.70**	0.56	15,678.4	**34,747.39**	0.70	6,350.1
G-H-05	16,138.50	-4.72	51.2	15,895.94	-3.14	799.6	**15,411.82**	0.00	754.2	15,411.82	15,714.24	-1.96	1,132.2	15,751.18	-2.20	444.4
G-H-06	20,329.04	-2.37	100.5	20,381.35	-2.63	1,356.9	**19,859.30**	0.00	1,954.1	19,859.30	20,116.24	-1.29	2,997.2	20,204.73	-1.74	1,149.5
G-H-07	24,184.83	-3.00	160.8	23,915.77	-1.85	2,262.5	**23,481.28**	0.00	9,167.3	23,481.28	23,751.16	-1.15	6,112.2	23,716.72	-1.00	2,527.9
G-H-08	27,710.66	-1.37	258.4	27,521.28	-0.68	3,408.2	**27,334.84**	0.00	18,625.2	27,334.84	27,482.48	-0.54	11,921.0	27,433.18	-0.36	4,632.0
G-H-09	1,346.03	-1.26	81.4	1,331.11	-0.14	592.7	**1,329.27**	0.00	1,829.3	1,329.27	1,373.90	-3.36	471.5	1,371.14	-3.15	167.0
G-H-10	1,575.82	-1.34	133.2	**1,554.96**	0.00	1,087.1	1,555.59	-0.04	1,564.3	1,554.96	1,611.09	-3.61	1,232.9	1,604.15	-3.16	489.4
G-H-11	2,218.91	-1.26	214.0	**2,191.23**	0.00	1,445.5	2,195.83	-0.21	3,207.9	2,191.23	2,246.13	-2.51	2,001.0	2,244.60	-2.44	782.4
G-H-12	2,510.07	-1.09	297.0	2,535.00	-2.10	2,108.3	**2,482.92**	0.00	4,224.0	2,482.92	2,554.43	-2.88	4,529.1	2,559.52	-3.09	1,689.7
G-H-13	2,253.45	-0.97	73.3	**2,231.88**	0.00	405.8	2,237.38	-0.25	1,801.3	2,231.88	2,307.28	-3.38	101.6	2,300.13	-3.06	35.0
G-H-14	2,711.81	-1.01	124.6	2,685.51	-0.03	630.3	**2,684.70**	0.00	1,042.9	2,684.70	2,752.90	-2.54	221.0	2,767.00	-3.07	74.9
G-H-15	3,156.93	-1.07	189.5	**3,123.60**	0.00	976.6	3,127.33	-0.12	2,111.2	3,123.60	3,222.33	-3.16	462.7	3,237.13	-3.63	151.7
G-H-16	3,649.09	-0.75	278.5	3,853.21	-6.39	1,571.2	**3,621.85**	0.00	6,217.4	3,621.85	3,768.94	-4.06	764.5	3,770.16	-4.09	254.3
G-H-17	1,705.48	-2.49	76.2	1,674.91	-0.65	396.9	**1,664.08**	0.00	522.6	1,664.08	1,680.94	-1.01	100.1	1,685.81	-1.31	33.1
G-H-18	2,759.99	-1.89	129.9	2,722.32	-0.50	617.7	**2,708.73**	0.00	1,132.5	2,708.73	2,730.16	-0.79	152.7	2,727.52	-0.69	49.3
G-H-19	3,517.48	-2.15	189.2	3,445.85	-0.07	841.8	**3,443.59**	0.00	1,627.1	3,443.59	**3,439.96**	0.11	221.7	3,449.35	-0.17	77.0
G-H-20	4,413.82	-2.49	273.3	**4,306.53**	0.00	1,392.2	4,314.16	-0.18	2,969.8	4,306.53	4,312.65	-0.14	275.7	4,334.54	-0.65	97.7
		-1.73	165.84		-1.13	1408.76		-0.05	4048.77			-1.62	2965.6		-1.70	1165.5

3 Distribution Districting – the case of in-night express services

Author: Christian Brabänder

This chapter is the peer-reviewed and published version of Brabänder, Christian "Distribution Districting – the case of in-night express services", in: Journal of Service Management Research (SMR), 2018, Vol. 2, Issue 4: pp. 36-52, DOI: 10.15358/2511-8676-2018-4-36. [37][38]

Abstract

Distribution districting is the problem of partitioning a logistics service area into districts. A good districting plan is characterized by its properties of contiguity, compactness, and balance. Herein, [an] integer program of the distribution districting problem is proposed. Compactness serves as the objective, since compact districts have greater operational routing flexibility. Balance of the estimated travel time per district is incorporated in the constraints of the model. Contiguity is modelled using a net flow formulation. The applicability is demonstrated at the case of in-night express at Kiessling Spedition. In express services, districting is of major importance since the feasibility and productivity of tours, bundling of shipments, daily routing efficiency, and satisfaction of customers and drivers are depending on a good districting plan. This paper contributes to the body of districting literature through a practical application of a quantitative model and thus bridges the gap between academic literature and practice.

Keywords

Districting, in-night express, distribution, service area, Mixed Integer Programming

[37] This paper is written and published in British English. Therefore, the orthography in section 3 deviates from the other sections.
[38] In some details marked by square brackets, this chapter deviates from the original publication.

3.1 Introduction

Logistics service providers (LSP) operate distribution networks. In the distribution, they exploit "economies of transportation" from bundling shipments from a multitude of consigners. LSPs use districting to partition their service area into fixed tour areas, in which the workload of every tour area is feasible and more or less balanced. Many heterogeneous consigners frequently feed shipments into the distribution system with varying quantity and destinations.

Every LSP needs to evaluate and re-optimize the districting plan from time to time. The reason is a change in the underlying shipment structure. On the one hand, large consigners entering a system require additional capacities in terms of delivery time and volume. On the other hand, large consigners leaving a system leave the LSP with idle capacities. This recurring problem baffles the LSP with a re-optimization problem: incumbent districts become imbalanced, infeasible, rule breaking, or run idle. As a result, the distribution capacities are adjusted to the changed shipment structure. Consequently, whenever the shipment structure of a distribution system changes, the operator must evaluate, adjust, and re-optimize the distribution districts (van Oudheusden et al. 1999, p. 525).

This paper addresses the districting problem in distribution services.

3.1.1 Distribution districting

Districting is the problem of partitioning a geographic area into smaller districts in such a way that the resulting districting plan is acceptable according to some planning criteria. Typically, there are some basic or elementary units in the area to be partitioned. Those basic units are grouped together into districts (Kalcsics 2015, p. 596). Kalcsics (2015, p. 597) categorized districting problems into four fields of application: political, sales territory, service, and distribution. This paper investigates distribution districting. Distribution districting aims to partition a logistics service area into districts such that one district is assigned to one operating vehicle for at least several months. The assigned vehicle serves all delivery destinations (stops) within the district's boundaries. The operative routing of stops within a district is performed on a daily basis (Feige and Klaus 2008; Dettenbach and Ubber 2015; Kalcsics 2015). Mid-term stability is the main reason why districting is deployed in the first place. There is uncertainty both in the number and in the dispersion of delivery locations.

Many authors propose to solve this issue through daily optimization of the vehicle routing problem, which assigns stops to vehicles and optimizes the routes of vehicles simultaneously (e.g. Golden and Assad 1988; Wong 2008; Cattaruzza et al. 2017). This approach aims at exploiting every potential saving and requires a high degree of operational flexibility at the cost of mid-term stability. To account for stability this approach is extended by driver learning (Kunkel and Schwind 2012; Schneider 2016).

Districting is a tactical planning problem that provides stability in the distribution system, and thus generates several advantages: drivers' familiarity with their districts, roads, and customers increases. As a result, the drivers' productivity increases through less detours, better ability to improvise when faced with exceptions like closed roads, and less time needed to search drop-off locations like gates or offices. Stability in the distribution reduces overhead costs, since processes are more standardised. The interface between the driver and the customer is smoother because they get familiar with one another (O'Brien 1975; Mole 1979; Wong and Beasley 1984; Savelsbergh and Goetschalckx 1995; Wong 2008). The major disadvantage of districting compared to daily new routing is a loss of operational routing flexibility. As a result, even good districts may perform poorly on some days due to unfavourable stop locations or stochastically unbalanced workload.

From a processual point of view, not every LSP can exploit routing flexibility but requires stability. For example, in-night express service providers have keys to unlock customers' depots for unattended "receiptless" delivery. For legal reasons, the provider must assign every key to one driver. The process of swapping keys, documenting the transfer of liability, and exchanging the necessary knowledge (e. g. which key fits to where? How does one get there?) is time-consuming and error-prone. As a result, in-night service providers use distribution districting for reasons of processual stability and standardisation.

3.1.2 Outline

The goal of this paper is to present [an …] Integer Program (IP) formulation for the distribution districting problem (DDP). Districting is subject to many constrains and requires a careful trade-off between several attributes and properties. This requires a substantial amount of rationalisation and time. Therefore, it is believed that [an IP] is a powerful model to generate solutions to the DDP. The research questions are: How to model the distribution districting problem using [an IP]? Do solutions for the DDP from an [IP] model perform well in practice?

In order to formulate the [IP], conceptual considerations about desirable properties of districts are collected and literature about districting having a focus on distribution is reviewed in section 2. In section 3, an [IP] formulation for the DDP is presented. Section 4 introduces the case of Kiessling-Spedition, an in-night express provider. This case study describes the real-world problem of a major consigner leaving the in-night distribution network. In section 5, using real-world data from the Kiessling case, the model is solved and the solution is evaluated. For the sake of comparability, two more solutions (an adapted savings and a practitioner's solution) are evaluated and compared. It turns out that the [IP] solution performs really well. However, also the practitioner's solution is doing almost as good as the [IP]. Section 6 discusses implications and further research topics.

3.2 Literature review

In this section, the theoretical foundations of distribution districting are outlined. Literature is reviewed in order to gain helpful insights to formulate an optimization model for the DDP. Subsection 3.2.1 collects conceptual considerations from literature and subsection 3.2.2 gives an overview of approaches to formulate a districting model.

3.2.1 Conceptual considerations for distribution districting

The long-term distribution planning is about the physical structure of the distribution network, whereas the mid-term distribution planning is about the line-hauling between hubs and terminals and districting. The daily truck loads and routes are planned in the short-term. Districting bridges the planning dimensions of long-term planning of the physical distribution system and the daily routing (van Oudheusden et al. 1999). The three planning dimensions are neither independent nor strictly hierarchical because the estimated daily routes need to be considered in location decision-making and districting. However, districting is usually done after locations have been fixed and before routes can be scheduled. Therefore, districting is an intermediate step in distribution planning (van Oudheusden et al. 1999, pp. 523–524). When planning the distribution districts, one has to follow some criteria. A complete and definite list of criteria is hard to pin down. The repeatedly considered properties of good districts are compactness, balance, and contiguity (Shirabe 2009; Kalcsics 2015).

3.2.1.1 Compactness

According to van Oudheusden et al. (1999), good districts are flexible with respect to future routing. Herein, the term flexibility means that, despite the demand and location uncertainty in daily routing, the districts facilitate robust routing without increasing neither time nor cost per shipment disproportionally (Upton 1994). In literature, routing flexibility is discussed in terms of compactness of districts. Compactness is pointed out as an essential property of districts in order to minimize daily travel times. Daskin (2010, p. 246) calls this property "convexity" synonymously. Basically, it means to design districts of approximately round or square shape (Shirabe 2009, p. 1054). It remains disputable how to measure and manage compactness. Many models require linear measures, others may also deal with non-linear ones. Among others, Niemi et al. (1990) outline some alternative measures. Furthermore, it remains case-specific whether compactness is the objective or a constraint of districting models.

3.2.1.2 Balance

The identification of imbalanced workload, feasibility of operations due to some constraints, and fairness of drivers' wages are discussed in literature. Activity measures are indicators of how much corporate activity takes place in a basic unit. Districts are balanced if they have a comparable level with respect to all relevant activity measures. The need to balance districts is widely recognized in literature. This criterion is synonymously named equilibrium or integrity. For example, Daskin (2010, p. 248) proposes to balance the total demand assigned to a district. Schmidthöfer (2004) contributes to distribution districting by balancing the estimated driving time. Other activity measures are area, stops (Lei et al. 2012), colli, total weight (Gelders and Cattrysse 1991), frequency of demand, and value per stop (Fleischmann and Paraschis 1988), total demand (Daskin 2010), population (Rodrigues and Ferreira 2015). Besides the expected workload, the uncertainty of workload is another activity measure of interest. Some basic units have a rather constant total demand and are hence certain. Other basic units have variable total demand and demand locations. These different degrees of uncertainty could be balanced across districts. Haugland et al. (2007) have considered stochastic demands in the context of stochastic districting. Many districting models employ the balance of the most important measure as the objective. Others treat balance as a constraint (Kalcsics 2015).

3.2.1.3 Contiguity

Contiguity is the third essential property of districts. This property reduces the daily travel times because drivers do not need to exit their district after they entered it once (Kalcsics 2015, p. 606).

van Oudheusden et al. (1999, p. 525) point out that good districts do not overlap and routes should not cross one another. This means, there should not be any enclaves in a solution because enclaves form districts of discontinuous parts. Shirabe (2009, p. 1054) points out that, while it is easy to directly recognize whether a district is contiguous or not, contiguity is often excluded from districting and it is implicitly assumed that compact districts are likely to be contiguous. Contiguity is always treated as a constraint.

3.2.2 Approaches to distribution districting

Fleischmann and Paraschis (1988) apply a transportation-location procedure for districting in the case of a large West-German manufacturer of consumer goods. Schmidthöfer (2004) develops a comprehensive optimization model for distribution districting based on a location-allocation-approach. He highlights the development of an advanced travel time estimator. The estimated travel time is an important activity measure in the DDP. Schmidthöfer's estimator builds upon the research of Beardwood et al. (1959), Daganzo (1984), and Blumenfeld and Beckmann (1985). He develops an objective function that minimizes the approximated travel distance over all districts. It seems like state-of-the-art to adopt the estimator from Beardwood et al. (1959). Haugland et al. (2007) design districts for the vehicle routing problem when demand is stochastic. They apply a two-phase procedure, first designing districts for unknown demand, and second routing vehicles when demand is known. They apply a tabu search heuristic to solve the problem. Lei et al. (2012) also use such a two-phase procedure for a combined districting and routing problem. They first use a large neighbourhood search heuristic to generate districts and then estimate the expected routing costs adopting the Beardwood et al. formula. Carlsson (2012) presents an algorithm to balance vehicles' workloads for independent and identically distributed customers within a polygon, using geodesic cuts.

The [IP] formulation by Shirabe (2009) contributes to districting research through enforcing contiguity by using a net flow approach. He outlines three different [IP] formulations based on this idea, using different objective functions which optimize either compactness or balance. Other approaches to enforce contiguity are graph-based models like Voronoi diagrams (Novaes et al. 2009) and the formulation inspired by sub-tour elimination constraints (Drexl and Haase 1999, p. 1310) as known from travelling salesman problems. Daskin (2010, pp. 244–262) formulates an [IP] model which minimizes the differences between the maximal and the minimal total demand assigned to a district. This formulation does not consider contiguity explicitly, but tries to design contingent districts through compactness measures. Dettenbach and Ubber (2015) present the problem of handling disruptions in distribution districts. They propose the concept of a back-up district.

Kalcsics (2015) presents an overview of districting applications, criteria, constraints, and models. Refer to this excellent summary for details on other applications than distribution. Franceschetti et al. (2017) summarize estimators in freight management. Among others, they summarize estimators used in distribution.

Districting is in theory a well understood topic that received a lot of attention from different disciplines (Kalcsics 2015). In the rich corpus of districting literature, there are only few real-world applications. However, it remains unclear whether compactness or balance should be used in the objective function and how to measure the criteria of compactness, balance (activity measures), and contiguity. For this reason, this paper contributes a case study using real-world data and an [IP] model specifically modelled for application in a distribution context. The [IP] approach by Shirabe (2009) for districting serves as a blue-print to formulate an [IP] model for the DDP. From Schmidthöfer's model the estimator of driving time is adopted.

3.3 Mixed Integer Program

For this paper, [an IP] model of the DDP is formulated. The general districting model by Shirabe (2009) is adapted to the DDP in three ways: first, the estimator of travel time is adopted from Schmidthöfer (2004) as the most important activity measure. The travel time is balanced between districts. Second, the time to approach a district is incorporated into the balancing constraints using estimates of the travel time starting at the terminal and approaching the centre of a zip code. Third, the minimal sum of weighted net flows is a new measure of compactness in the objective function. Optimizing compactness has the benefit of better daily routing flexibility and produces robust solutions for unexpected high number of shipments.

3.3.1 Activity measure: estimated travel time

The question how long a distribution vehicle will drive within an area is not trivial to answer. It depends on many determinants: size of the area, number of stops in the area, dispersion of stops across the area, possible velocity, density and quality of the road network, and occurrences of natural obstacles like rivers. Since the number and location of stops per night is uncertain, an estimate of the travel time is adopted from Schmidthöfer (2004). Lei et al. (2012, p. 70) apply the same approach as an activity measure in order to estimate the expected cost of routing per district. The foundation of the estimators in both papers is the formula of Beardwood et al. (1959), which establishes the approximate relation of Euclidean travelling salesman problem (TSP) tour length, area, and number of vertices.

$$\bar{d} = \beta \cdot \sqrt{n \cdot a} \qquad (1)$$

The Beardwood et al. TSP-formula (1) states that estimated [tour] length $\bar{d}$ is proportional to the square root of the product of the number of vertices visited n and the area of size a .

The estimated values for the constant β may vary depending on the shape of the area (Daganzo 1984), the distance metric in-between vertices, the number of stops (Lei et al. 2012), and the dispersion of vertices. Applegate et al. (2011, pp. 493–500) perform extensive testing to estimate the value of β using random Euclidean distances. They find a downward curve for β as n increases. Schmidthöfer (2004, p. 132) derives the following estimator of travel time $\bar{t_i}$ per zip code i :

$$\bar{t_i} = \frac{\beta \cdot \sqrt{(n_i - 1) \cdot a_i} + \lambda \cdot \rho_i^{-1/2}}{v} \qquad (2)$$

In formula (2) the coefficient n_i is the expected number of stops on the tour, a_i is the size of the area, v is the average velocity in regional traffic, ρ_i is the density of stops n_i / a_i, and β and λ are constants. The constant λ is introduced as a correction for the unknown approaching direction (where does the vehicle come from?) and the mean distance between stops of different zip codes (how far is it from the last stop of a zip code to the first stop of the subsequent zip code?) (Schmidthöfer 2004, pp. 108–110). The number of stops is reduced by 1 in (2). While this correction is theoretically correct for high values of n_i (there is no TSP tour, as the vehicle exits the zip code area after the last stop), it is highly problematic in the case of very low expected number of stops. In the case study, there are some zip codes having like 0.5 stops per night on average. For this reason, this correction is herein neglected in this case study.

The estimator is augmented by the stopping time. Stopping time is estimated using a linear model:

$$\overline{s_i} = n_i \cdot \overline{s^f} + c_i \cdot \overline{s^v} \tag{3}$$

The estimated stopping time (3) has a fixed component $\overline{s^f}$ per stop and a variable component $\overline{s^v}$ per delivered colli. The measure c_i is the number of delivered colli per zip code area. For the remainder of this paper $\overline{s_i}$ is added to $\overline{t_i}$ for the sake of readability and (4) is applied hereafter.

$$\overline{t_i} = \frac{\beta \cdot \sqrt{n_i \cdot a_i} + \lambda \cdot \rho_i^{-1/2}}{v} + n_i \cdot \overline{s^f} + c_i \cdot \overline{s^v} \tag{4}$$

Other activity measures that are important in distribution are the number of colli, volume, and weight of the total vehicle load. The vehicles used in regional distribution are usually vans having a payload of max. 1.4 tons or box trucks having a payload of max. 3.5 tons. In practice, it is important to discuss which constraints are most binding and incorporate those first. In the case of in-night express, the estimated travel time is the top priority, since the guaranteed service time is usually the only binding limitation. From a customer's perspective, he pays a great premium for the guaranteed time of delivery because own subsequent processes are dependent on punctual delivery. Nonetheless, if the necessary input data is available, other activity measures $q \in Q$ can be included and balanced likewise.

3.3.2 Neighbourhood graph

The net flow paradigm adopted from Shirabe (2009) requires an underlying graph of vertices and arcs. The natural representation of the physical service area is a geographic map consisting of the zip codes that represent the basic units. Every basic unit is a contingent polygon. This map needs to be transformed into a neighbourhood graph. Every basic unit is represented by one vertex. Every vertex has some activity measures Q. A vertex's coordinates are the mean coordinates of all the stops within the represented zip code. Arcs represent neighbourhood relationships between basic units. Consequently, if two polygons are adjacent to one another, there are two arcs pointing from one basic unit's vertex to the other's vertex and vice versa. Every arc has a weight assigned representing the driving time between the neighbouring vertices. The set of vertices I and the set of arcs A are defined as the neighbourhood graph (I, A) of the service area.

3.3.3 [IP] model

In the DDP model, the following coefficients and decision variables are defined:

Sets and indices

I	Set of zip codes' vertices/basic units $i, j \in I$
A	Set of arcs $(i, j) \in A$ from zip code i to j
K	Set of districts $k \in K$ where $\|K\| = p$ is the given number of districts to be designed
Q	The set of $q \in Q$ activity measures of the zip codes' vertices

Decision variables

f_{ijk}	Integer decision variable of positive quantity flowing from i to j in district k
x_{ik}	Boolean decision variable, 1 if zip code i is assigned to district k , 0 otherwise
w_{ik}	Boolean decision variable, 1 if zip code i is chosen to be the centre of district k , 0 otherwise

Coefficients

$\overline{t_i}$	Estimator of the travel time within zip code i
t_{ij}	Estimator of the travel time from the centre of i to the centre of j
g_i	Estimator of the travel time from the terminal to the centre of zip code i
b_i^q	The value of the q^{th} activity measure of zip code i
B^q	The target value per district of the q^{th} activity measure
T	The target value per district of the total available time per day/night
δ^{t+}, δ^{t-}	Upside and downside tolerance of the total estimated time per district
δ^{q+}, δ^{q-}	Upside and downside tolerance of the q^{th} activity measure per district
m	Upper bound for the number of zip codes $i \in I$ to be assigned to district $k \in K$

The DDP is modelled using the following formulation:

Min $$\sum_{(i,j)\in A} f_{ijk} \cdot t_{ij} \tag{5}$$

s. t. $$\sum_k x_{ik} = 1 \qquad i \in I \tag{6}$$

$$\sum_i w_{ik} = 1 \qquad \forall k \in K \tag{7}$$

$$\sum_{(i,j)\in A} f_{ijk} - \sum_{(j,i)\in A} f_{jik} \geq x_{ik} - w_{ik} \cdot m \qquad \forall k \in K, i \in I \tag{8}$$

$$\sum_{(j,i)\in A} f_{jik} \leq (m-1) \cdot x_{ik} \qquad \forall k \in K, i \in I \tag{9}$$

$$\sum_i \overline{t_i} \cdot x_{ik} + \sum_j g_j \cdot w_{jk} \geq \left(1-\delta^{t-}\right) \cdot T \qquad \forall k \in K \tag{10}$$

$$\sum_i \bar{t}_i \cdot x_{ik} + \sum_j g_j \cdot w_{jk} \leq \left(1+\delta^{t+}\right) \cdot T \qquad \forall k \in K \qquad (11)$$

$$\sum_i x_{ik} \cdot b_i^q \geq \left(1-\delta^{q-}\right) \cdot B^q \qquad \forall k \in K, q \in Q \qquad (12)$$

$$\sum_i x_{ik} \cdot b_i^q \leq \left(1+\delta^{q+}\right) \cdot B^q \qquad \forall k \in K, q \in Q \qquad (13)$$

$$f_{ijk} \in \mathbb{N}_0^+ \qquad \forall (i,j) \in A, k \in K \qquad (14)$$

$$x_{ik}, w_{ik} \in \{0,1\} \qquad \forall i \in I, k \in K \qquad (15)$$

The objective function (5) is the sum of weighted flow volumes. Minimizing the network flow is a proposed measure for maximizing compactness. Intuitively, this can be visualized using the example of five vertices in Fig. 3.1. Every vertex has a supply of 1 and every district centre is a sink consuming at most m units. Long chains produce huge flows as they approach the district's centre. Short flows having only few hops are favourable compared to long chains. Weighting flows by their geographic length further improves compactness, as very long arcs are less favourable in terms of compactness. Note that the objective function (5) is a proxy for compactness but has no other logistics interpretation. Other measures of compactness from Shirabe (2009) have been tested, but did not produce visually persuasive and practical solutions. The constraints (6) ensure that every basic unit i is assigned to exactly one district k. Constraints (7) chose one basic unit i to be the centre of district k. The centre has no real-world importance for the districting plan. Centres are merely used as the sinks of the flows which enforces contiguity, and maximize compactness of the districts. Additionally, centres are used to estimate the approach time from the terminal to the district. The constraints (8) and (9) are the net flow constraints that ensure continuity and are adopted from the problem 2 by Shirabe (2009, p. 1060). The idea is that every basic unit has a supply of 1 and all the supply flows through the graph (I,A) into the assigned district's centre/sink. The next four constraints (10)-(13) implement the balancing of activity measures among the districts. First, (10) and (11) balance the distribution time. The estimated travel time per district, including the time for stopping and making a delivery, plus the approach time from the terminal to a district's centre, must fall into an interval of $\left[T-\delta^{t-}, T+\delta^{t+}\right]$ around the target travel time T per night. Note, this estimation of travel time does not consider a return journey. Other applications may incorporate the way back. The rational herein is the focus on the punctual delivery, but not on timely return. Subcontractors return to their private homes after making the last delivery. The other activity measures $q \in Q$ for all districts $k \in K$ are balanced by constraints (12) and (13). For every $q \in Q$ there is a target value B^q enforced, having an upside and a downside tolerance. Finally, (14) makes sure the flow decision variables f_{ijk} are integer and positive or 0, and (15) enforces the other decision variables x_{ik} and w_{ik} to equal 1 or 0.

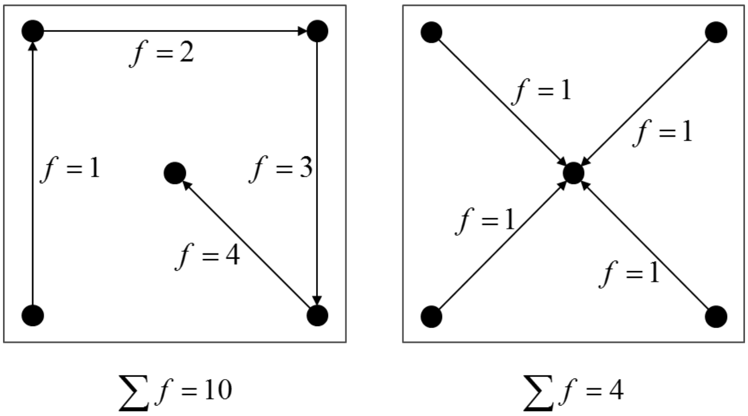


Fig. 3.1 Minimum net flow as a measure of compactness

3.4 Case Study

3.4.1 Express services

Courier, express, and parcel (CEP) service providers bundle […] shipments in their open hub-and-spoke network in order to maximize the […] utilization (Bretzke 2008, pp. 213–218). They bundle shipments on the collection, the line-haul, and the distribution leg. A very large number of different customers and a very high degree of process standardisation characterizes CEP providers. Therefore, the service is very fast, high frequent, and has low costs per shipment, but the service does not consider individual customer needs as staff do not have customer-specific knowledge and the equipment is highly standardised. Among CEP providers, express and courier providers offer guaranteed delivery at a promised time. Unlike couriers, express service providers run dedicated open networks to utilize "economies of transportation". The shipment size (weight and volume) is heterogeneous and the value density is relatively high compared to standard parcels, as the transport cost per shipment is higher due to the tough time constraints (Bretzke 2008, p. 217).

3.4.2 In-night express services

In-night express is a special case of express services. The fulfilment process of in-night service providers (INSP) works like that:

- collecting shipments until afternoon,
- transshipment at the terminal,
- line-hauling to the network's hub in the evening,
- transshipment at the hub before midnight,
- [line]-hauling [back] to the terminal,
- transshipment to the distribution vans, and
- finally distribution to the recipients until morning.

The special feature of INSP compared with conventional express providers is the so-called "receiptless" delivery at night-time: since there is nobody awake to accept the delivery and sign for it, there is no receipt issued. Instead, the driver scans the parcel, takes a photograph, and both a timestamp and the geo-coordinates of the delivery are saved. Usually, the delivery is placed in some kind of deposit box. Anything lockable can be a deposit box: mailboxes, garages, trunks, garden sheds, or storage depots. As a result, a driver may carry up to 300 different anonymous keys and […] needs customer-specific knowledge, where to deposit deliveries and how to get there.

3.4.3 Kiessling-Spedition [as an INSP]

Kiessling-Spedition (Kiessling) is a partner […] in a multinational in-night express network. Their terminal is located in Regenstauf (Eastern Bavaria, Germany). Besides in-night express services, they offer various logistics services, such as hazardous materials warehousing, contract logistics, procurement-logistics, and just-in-time concepts. Therefore, Kiessling is able to exploit synergy effects from utilizing its terminal hall both at day and night. This case study examines a problem of Kiessling's in-night express service. Kiessling's in-night line-hauling vans coming from the network's central hub (being located in the state of Hesse) arrive at the terminal around 2am. Usually, the distribution vans depart until 02:45am and perform the distribution to the recipients until 7 or 8am. The process is very time critical because delivery until 8am is guaranteed. There is a penalty enforced by the network for delayed deliveries. Subcontractors deploying their own vans perform the distribution. They operate utility vans, as this vehicle class has no specific speed limit in Germany. For the distribution, Kiessling uses districts. Every district is a set of zip codes, which are historically grown, negotiated with the subcontractors, and manually changed in the past.

3.4.4 Kiessling's need for new in-night distribution districts

In 2017, Kiessling faces the problem of losing its largest consigner, who is a large wholesaler of auto parts. Until then, this consigner fed very high volumes into the in-night network. A first analysis of historical data has revealed an expected loss in total shipments to Kiessling's service area of 40%. The number of stops is expected to drop by 26%. Hence, the stop-factor, being the ratio of shipments and stops, decreases by the factor $\frac{1-0.4}{1-0.26}=0.81$. As a result, the productivity in terms of delivered shipments per tour also decreases. Since INSP exploit synergies from fully utilized tours, the number of tours needs an adjustment. As there is one distribution district per tour planned, the managerial planning involves the re-design of the districting plan.

The case of Kiessling's in-night distribution districting serves as an example of the general problem of how to partition a service area into districts for distribution operations.

3.5 Application for Kiessling's in-night distribution

In the Kiessling case, the full service area is decomposed into five subareas. Every subarea is then districted into fixed tour areas. This decomposition has managerial reasons: first, for the sake of acceptance by the drivers, the new districting plan should have some similarity with the old one. Second, as the decision-makers at Kiessling have not used OR techniques for districting before, they aspire to develop a good feeling about this quantitative approach. Third, they want to test the newly designed districts in one subarea and later proceed with the others. This sequential approach permits learning throughout the project.

This section presents the results for one of these subareas, the northern part of Kiessling's service area. Ex-ante, there were four districts which were historically grown. Fig. 3.2 shows the old districting plan. Note, for the sake of viewer's orientation, the map displays larger towns. As one can easily see, there are already two issues within this plan:

- Tour 309 is not contingent
- Tour 803 is not compact as there is a finger-like shape of tour 311 reaching into the district

A decision-maker should address those issues from a districting perspective.

The exit of the largest consigner reduces the number of stops per night considerably. On average, the four tours lose 10.6 stops per night and tour, accounting for 27.5% of all stops in the northern subarea. Consequently, this area is now districted into three new districts and the cost of one tour is saved. As a result, the area per district, the mean distance in-between two arbitrary consecutive stops, and thus the expected travel distance per tour increase. The decision to eliminate one tour can be seen as deflating the metaphorical air buffer that was generated through the consigner's exit.

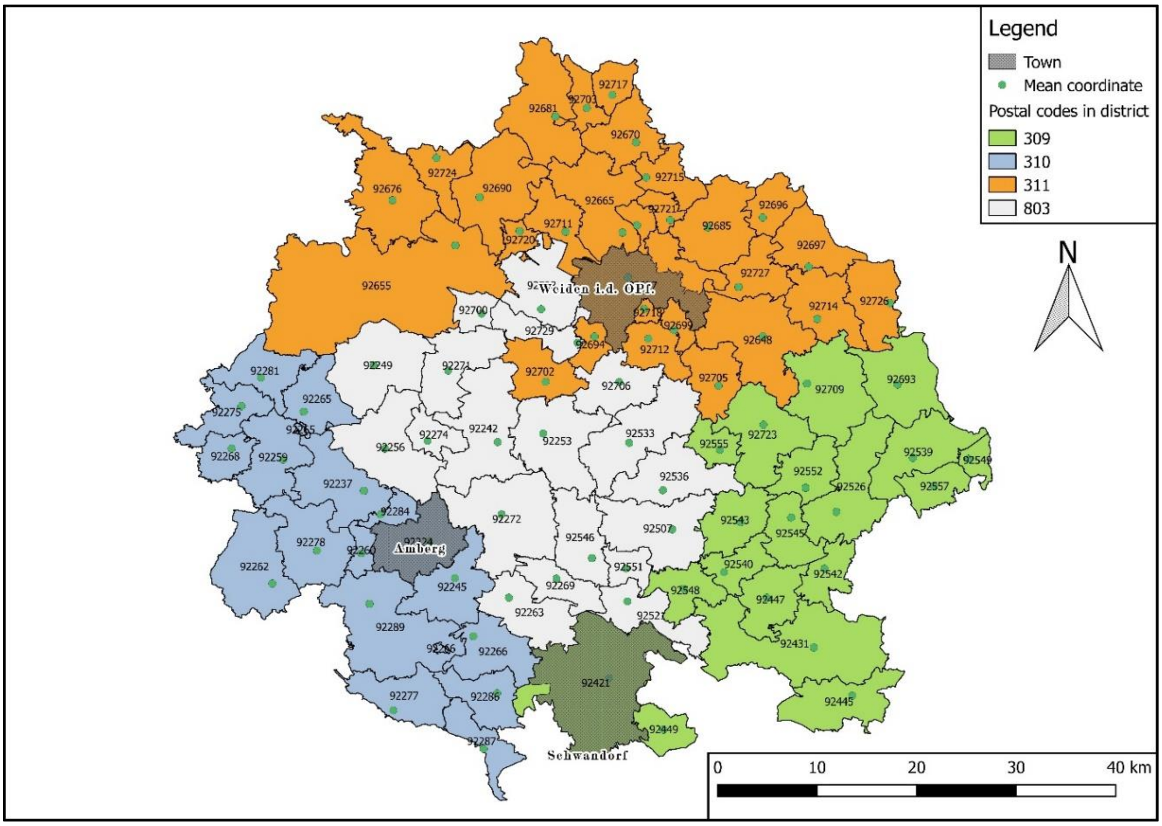


Fig. 3.2 Map of Kiessling's northern service area before the project

3.5.1 Data Input

3.5.1.1 Transactional records

Historic transactional records from Aug-Dec 2017 are extracted from Kiessling's Transport Management System (TMS). One record is one shipment. Every shipment has a date, a consigner, a recipient's address, the number of colli, the total weight, and the GPS coordinates from scanning the delivery. The records are aggregated on a stop level. A stop is defined as all the shipments delivered at the same date to the same coordinates. In order to identify identical coordinates, all coordinates within ten meters get a unique address-ID assigned. The stops are then categorized into "lost", "mixed", and "others". "Lost" stops are those receiving only shipments from the leaving consigner. Hence, only "mixed" and "others" stops are used in the DDP calculation.

3.5.1.2 Neighbourhood graph

For the neighbourhood graph, the geographic location of all 83 zip codes is required. For the project, the geographic information about zip codes is retrieved from Open Street Map (OSM)[39] and is cleaned from slivers, holes, and minor overlaps using QGIS[[40]] functionality. The shape file

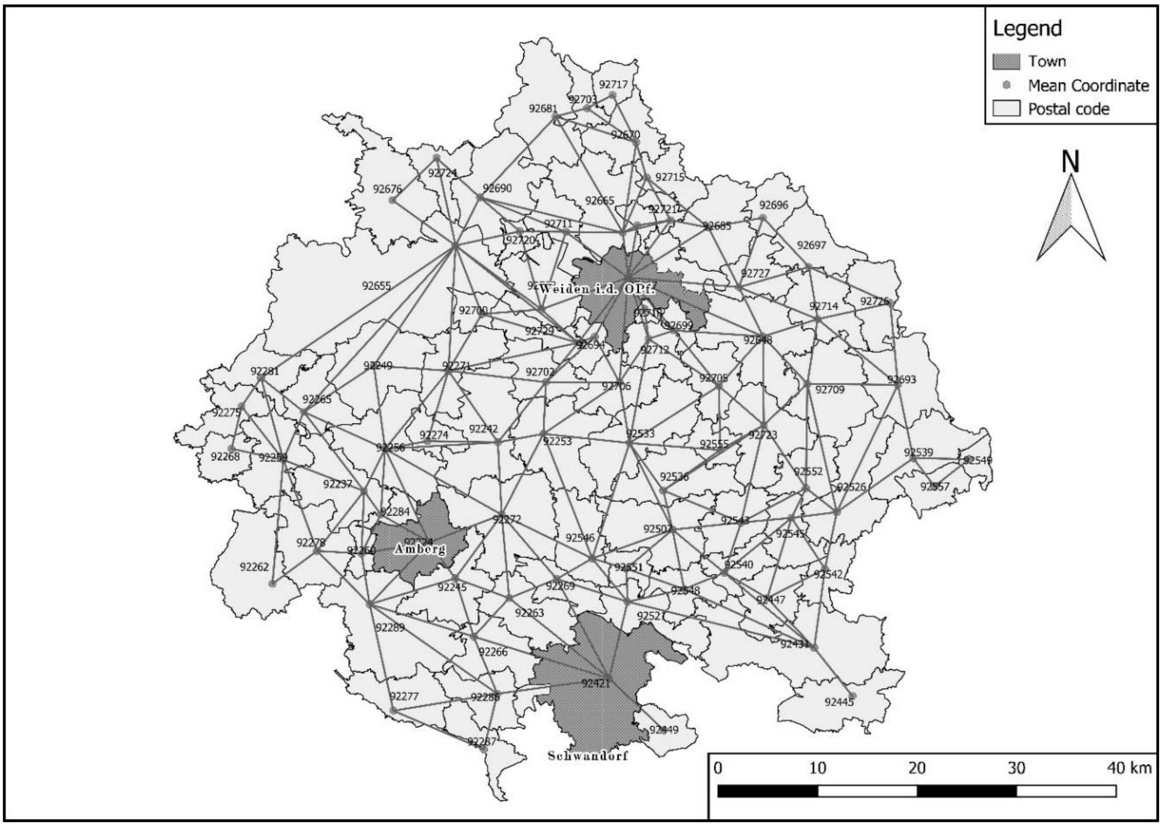


Fig. 3.3 Neighbourhood graph

is then browsed for neighbouring zip codes/polygons and a list of neighbours per zip code is extracted. *Fig. 3.3* depicts the neighbourhood graph. Every polygon is represented by one point within its boundaries. As many zip codes have clustered recipient addresses, the mean coordinate of all stops within a zip code in the sample has been chosen (green points). The travel time for every arc of the neighbourhood is retrieved from Google Maps via the Distance Matrix API[[41]].

---

[39] The polygon shape-files have been downloaded from the online resource https://www.suche-postleitzahl.org/

[40] QGIS is a freeware Geographical Information System (GIS). GIS are software that assist with collecting, editing, managing, reorganizing, analyzing, and presenting spatial data both visually and alphanumerically. More information is available online at https://qgis.org.

[41] The Distance Matrix API is a service of the Google Maps Platform that provides travel distance and travel time in a standardized format. An example of a valid URL to retrieve data from this API is: https://maps.googleapis.com/maps/api/distancematrix/json?origins=49.46247,12.426228&destinations=49.442689,11.765913&mode=car&language=de&units=metric&sensor=false&key={your_API_key}
By the time of the publication of this chapter, the use of this API was free. Nowadays, an active billing account is required.

3.5.1.3 Approach time

The time required to travel from the terminal in Regenstauf to a district is called the approach time. Neither returning to the terminal nor ending at some specific point (like the drivers' homes) is considered. It is just more important to finish the tours before 8am than minimizing the return journey's distance.

The data of the approach time is retrieved from the Google Maps Directions API. To this end, the coordinates of the terminal and the mean coordinates per zip code are used to receive realistic travel times based on navigation data.

3.5.1.4 IT landscape

The data and IT resources deployed in the case study are summarized in the landscape in Fig. 3.4. First, the historical TMS data is loaded, cleansed, aggregated, and augmented with geographical information. Second, the neighbourhood graph is generated from cleaned and augmented OSM data in QGIS. Third, navigation data from Google Maps is retrieved. The spreadsheets with the input data are then processed using the IBM ILOG Optimization Studio. There the [IP] is implemented and solved. The results are written in a spreadsheet and visualized on a map in QGIS (see Fig. 3.5).

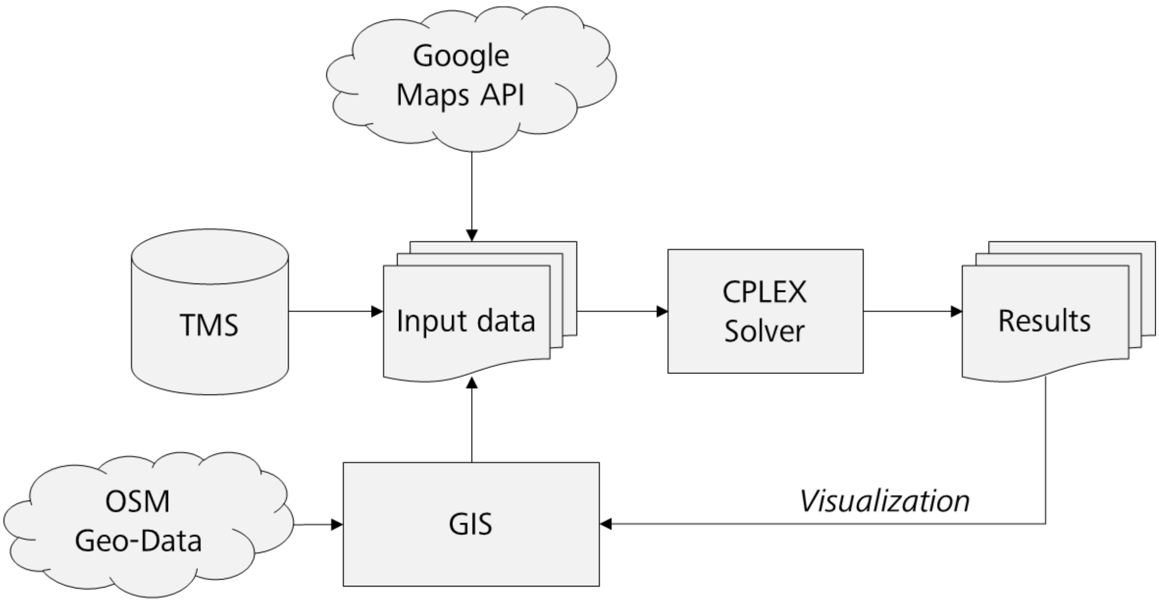


Fig. 3.4 IT landscape in the distribution districting case

3.5.2 Computational result

The DDP needs a carefully parameterized [IP]. There are several parameters in the [IP] model (5)-(15) which are set from experiments like the parameter m. Setting m too small makes the model infeasible, setting m too big increases the computation time. For the case study, $m = 33$ is set because this does not limit the solution space. In addition, the tolerance for the balance may be subject to further debate. For this case study, the 83 zip code areas have been partitioned into 3 districts. For the computation in the IBM ILOG Optimization Studio, the following parameter settings have been used:

$p = \|K\|$	3
T	Sum of estimated driving time of the basic units divided by $p = 3$
B^{area}	Total area of the northern service area divided by $p = 3$
δ^{t+}, δ^{t-}	10% tolerance
$\delta^{area+}, \delta^{area-}$	20% tolerance

The estimation of travel times within an area $\overline{t_i}$ was calculated using formula (4). The estimator is parameterized using the following setting:

β	0.765 (Schmidthöfer 2004, p. 160)
λ	1.15 (Schmidthöfer 2004, p. 160)
v	60 km/h
$\overline{s^f}$	2.5 min
$\overline{s^v}$	0 min

The value of velocity v being 60 km/h is actually quite high for regional distribution. However, using this assumption has resulted in good estimations in the in-night case. Testing in sub-section 3.5.4 will demonstrate that this value is a reasonable assumption. It may be justified by the fact that there is significantly less traffic at night-time and many traffic lights are turned off. Moreover, there are large rural areas and motorways in the service area with no or very high speed limits. In-night drivers reported even higher subjective estimates of their own driving speed. The fixed stopping time of 2.5 min. is an average over the course of one night. Since this parameter has a strong impact on the values of $\overline{t_i}$, the measured estimate is further validated using a bottom-up process model listing every single activity from applying to releasing the handbrake. The resulting bottom-up estimate for $\overline{s^f}$ is 140 sec. Including a little tolerance, 2.5 min is an accurate estimate. $\overline{s^v} = 0$ min is set due to a lack of data.

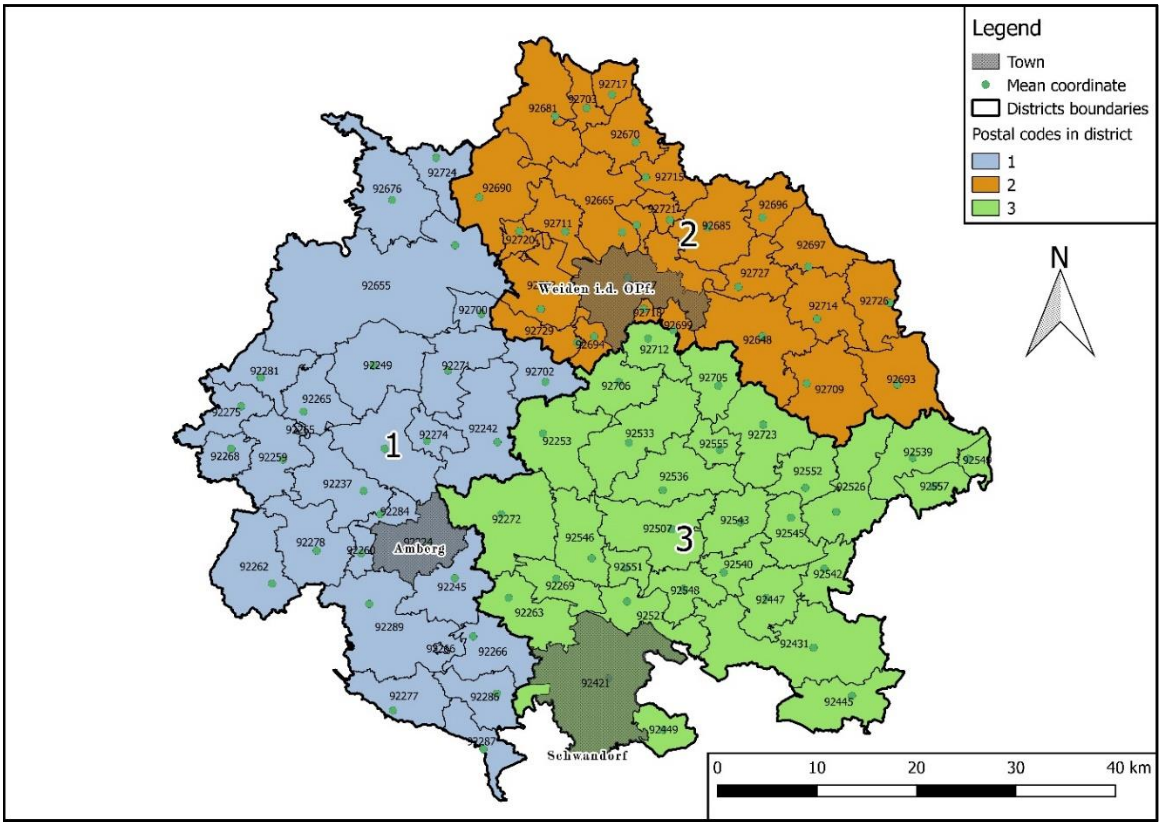


Fig. 3.5 Map of Kiessling's northern service area with 3 optimized distribution districts

The model is solved on an Intel Core i7 CPU at 2.8 GHz having 16 GB of RAM installed. The optimized assignment variables x_{ik} are visualized in Fig. 3.5. The minimized weighted network flow is 116,757. The total calculation time reported by the IBM ILOG Optimization studio is 265.06 seconds. The minimized solution of the network flow is visualized in Fig. 3.6. The labels on the arcs are the unweighted quantity flowing towards the centre.

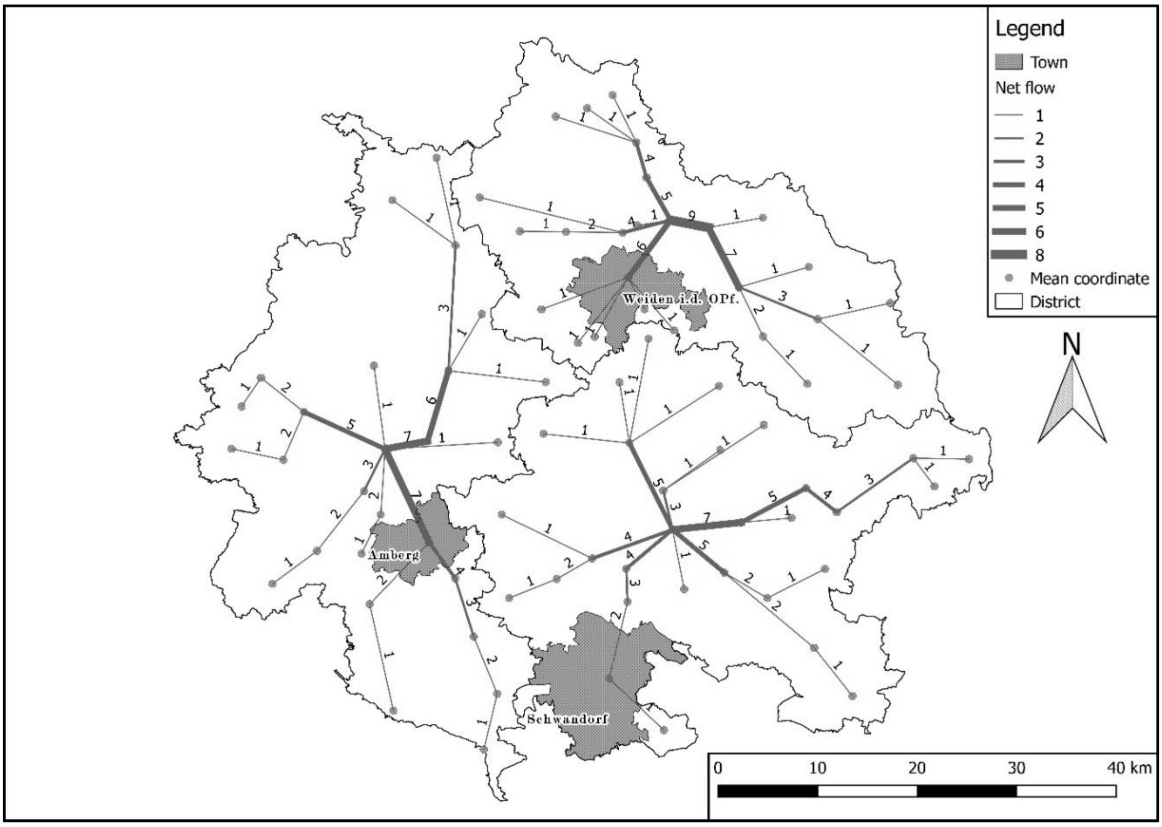


Fig. 3.6 Minimal network flow

Visually one can observe that the three districts are contingent and there are no obvious misdemeanours against the compactness criterion. The three largest towns in the service area are evenly spread over the districts. The terminal is located south of the service area. As a result, the approach time to district no. 2 is the longest and thus the area of this district is the smallest.

From a practitioner's view, the estimated duration per tour is the most relevant outcome. The driving time in a district is calculated from formula (4) with the districts area and the total number of expected stops per district. The approach time is the minimal driving time from the Regenstauf terminal to a basic unit's mean coordinate. Stopping time is the expected number of stops per night multiplied by $\overline{s^f} = 2.5$ minutes. Fig. 3.7 itemizes the components of the estimated tour duration in hours.

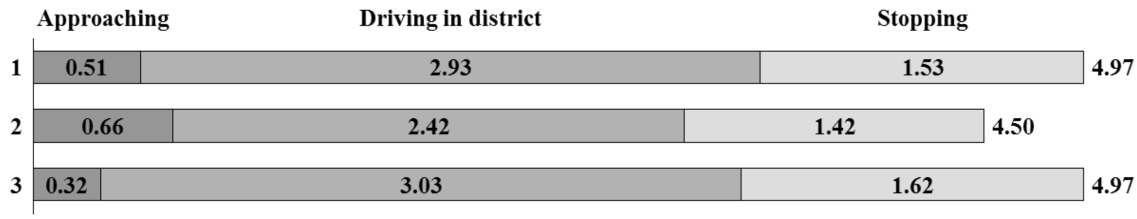


Fig. 3.7 Estimated tour duration [h] of optimized districts

As the three tours are almost balanced evenly and the total durations are expected to be feasible and smaller than the available time of 5.25 hours, this tour plan is recommended. However, in practice there are reservations regarding the districting plan and a few manual changes have been made. The prevalent practitioner's arguments in favour of manual changes are:

- Drivers' familiarity with some regions
- Difficult routing options due to the course of a river
- Differences between drivers' productivity/capabilities
- Subcontractors' additional, flexible capacities

3.5.3 Alternative solutions for comparison

In order to compare the quality of the [IP] solution from sub-section 3.5.2, two alternative solutions are proposed: first, an adapted savings algorithm is applied to design distribution districts, second, a "practitioner's solution" is designed from scratch following simple guidelines.

3.5.3.1 Adapted Savings Algorithm

Following the idea by Clarke and Wright (1964), the well-known savings algorithm is a first attempt to build tours. Whereas the original algorithm is meant to find tours through many points subject to limited vehicle capacities, districting partitions basic units, which are polygons, in the service area into districts. Therefore, the original algorithm is adapted to the DDP. The neighbourhood graph (*Fig. 3.3*) is used to represent the basic units and their relation. Equation (4) is applied to estimate the incumbent districts and to rank savings. The adapted savings algorithm works like shown in Table 3.1:

Table 3.1 Adapted savings algorithm

Initialization
▪ For each basic unit (zip code) initialize one district ▪ Evaluate the travel time using (4) plus the approach time from the terminal ▪ For each arc of the neighbourhood graph compute the savings s_{ij} from merging two adjacent districts i and j using equation (16)-(18): ○ Add the area of merged districts a_i and a_j ○ Add the stops of merged districts n_i and n_j ○ Apply the minimum of approach times from the terminal $$s_{ij} = \left(\overline{t_i} + g_i + \overline{t_j} + g_j\right) - \left(\overline{t_{i+j}} + g_{i+j}\right) \quad (16)$$ $$\overline{t_{i+j}} = \frac{\beta \cdot \sqrt{\left(n_i + n_j\right) \cdot \left(a_i + a_j\right)} + \lambda \cdot \left(\frac{n_i + n_j}{a_i + a_j}\right)^{-1/2}}{v} + \left(n_i + n_j\right) \cdot \overline{s^f} \quad (17)$$ $$g_{i+j} = \min\left\{g_i, g_j\right\} \quad (18)$$ ▪ Rank all savings s_{ij} in descending order
Iteration

- Evaluate if the greatest savings are feasible: $\overline{t_{i+j}} + g_{i+j} \leq 5.25h$ where $\overline{t_{i+j}}$ is the estimated driving time of merged districts i and j and g_{i+j} is the minimal driving time from the terminal to a basic unit of the merged district
    - If feasible: implement the best merge
    - If not feasible: forbid merge and go to the next greatest savings
- Update all savings of neighbouring districts
- Rank remaining savings in descending order

Abort

- Best remaining saving is $s_{ij} < 0$ or
- List of permitted savings is empty

The algorithm is implemented in Visual Basic for Application and runs using the same input data as described in sub-section 3.5.1 and parameter values in sub-section 3.5.2. The solution of the adapted savings algorithm is shown in Fig. 3.8.

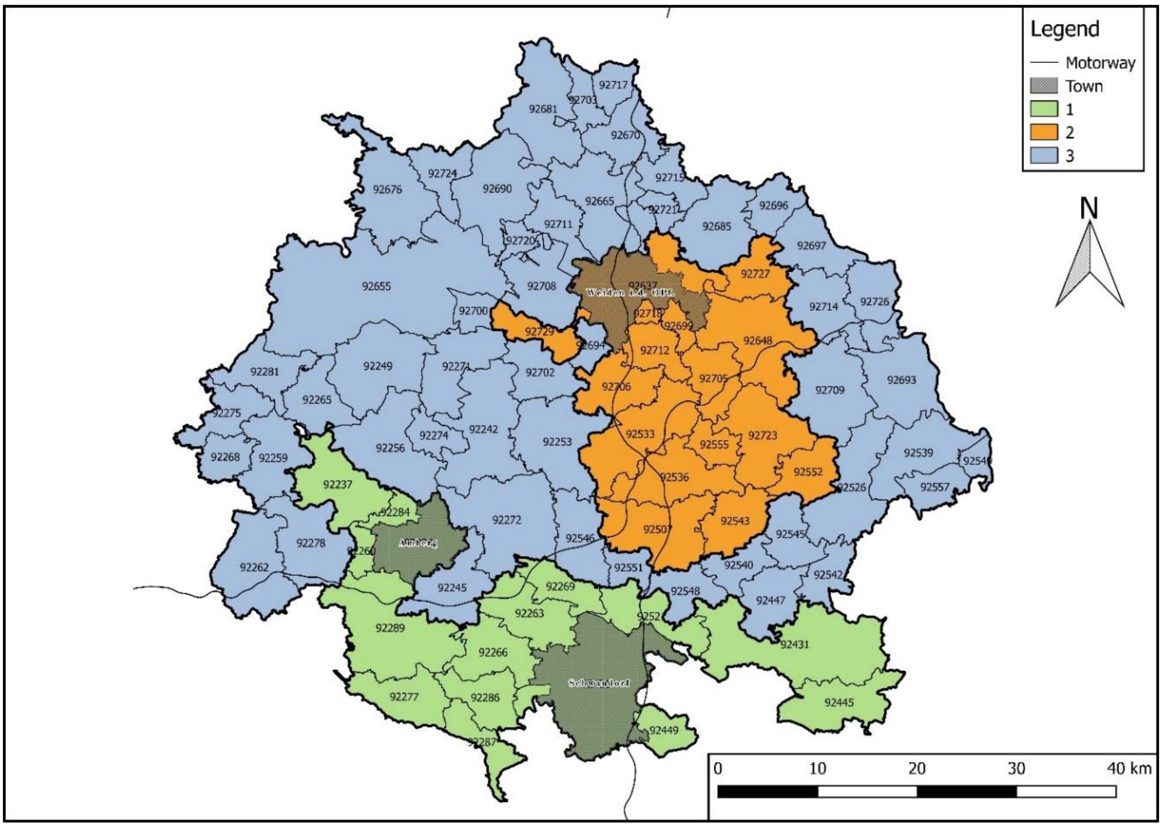


Fig. 3.8 Adapted Saving Algorithm solution

With respect to the contiguity, one can easily observe that the three districts are contingent. Regarding the compactness, this solution is not compact. There is an enclave (district 2) within district 3. Further, district 1 is not of "round" shape. Regarding the balance of the solution, there occurs a problem with the estimator (4): since the Beardwood et al. (1959) formula is designed for uniform distribution of points in a bounded area, it is not applicable to district 3. There are no stops of district 3 within the enclave. Therefore, the assumption of uniformly distributed points is not sufficiently met.

If the estimator is applied nonetheless, the estimated tour durations of the three districts including stopping and approaching are 4.48h, 4.07h, and 5.18h. As district 3 has a 27% greater estimated duration than district 1, the solution is not as balanced as the [IP] solution. Further, in district 1 there are the towns of Schwandorf and Amberg. In district 3 there is a great area of land, but less stops. Consequently, the driving time in district 3 is expected to be very long while the stopping time is comparatively small.

There is another drawback of the algorithm: it does not guarantee to minimize the number of districts. It may happen that there are further enclaves neighbouring only one large district which is already fully utilized. This issue does not occur here – notice it anyway.

3.5.3.2 A practitioner's approach

A "practitioner's approach" is a rule-based educated guess [performed by an experienced manager]. Regarding the districting problem at hand some simple rules are deployed:

- The three major towns (Weiden, Schwandorf, Amberg) must be allocated to the three districts.
- There is a motorway A93 leading from the terminal northwards into the territory. Every district must be within close proximity of the A93.
- There is a second motorway A6 crossing the A93 leading eastwards. This motorway should be on the border between districts, such that the western and eastern part of districts are connected along this motorway.
- The three properties contingency, compactness, and balance of area should be fulfilled visually.

The author and the responsible decision-maker at Kiessling have districted the service area using these guidelines. Fig. 3.9 shows the proposed solution.

The depicted solution is contingent. With respect to compactness, no serious violations can be detected. The proximity to both motorways is considered as well as possible. The three larger towns are allocated to the three districts.

3.5.4 Testing

The optimized distribution districts from the [IP] model and the alternative solutions are tested in a routing software tool employing real-time navigation data[42]. Vehicle routing starts at the terminal and computes the fastest (in contrast to the shortest) path through all the stops of one night ending at one of the stops.

[42] The software employed here is an in-house routing software from Kiessling. This software uses navigational data from the Graphhopper API for computation of the operational routes and visualized the computed tours in an online GIS.

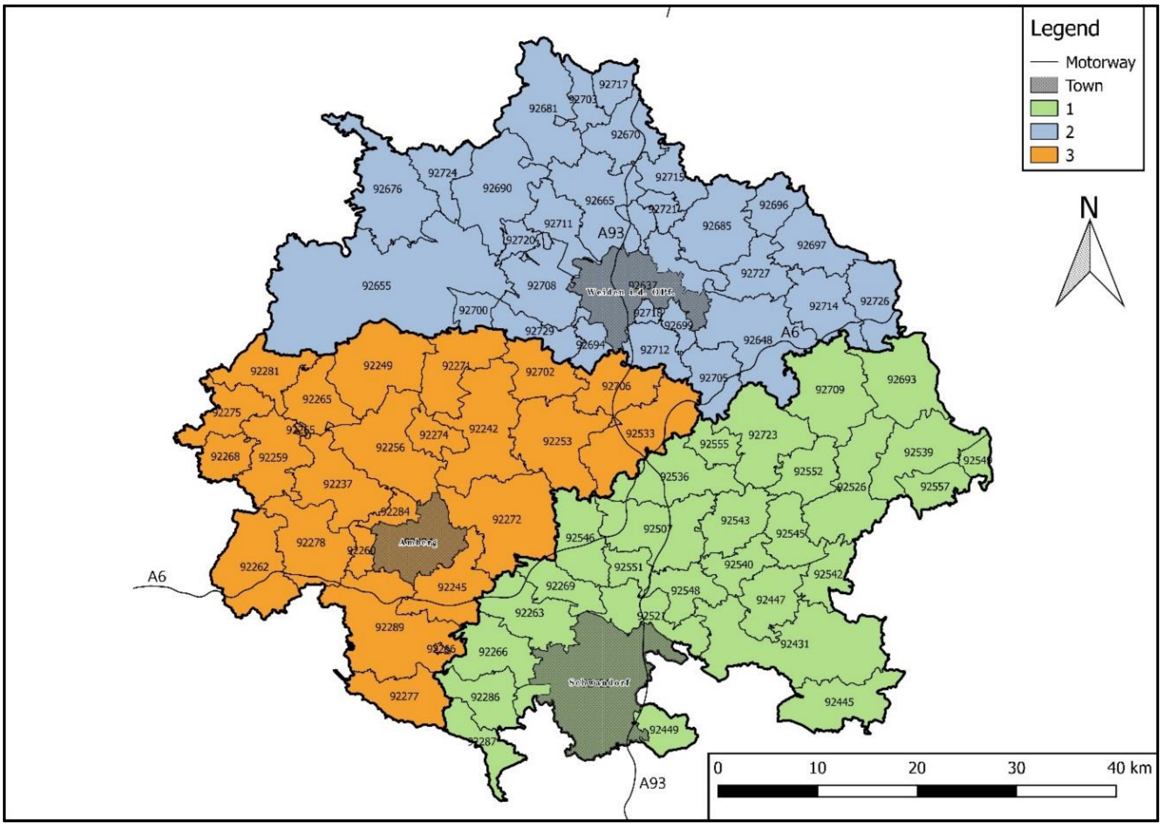


Fig. 3.9 A practitioner's solution to the case

3.5.4.1 Testing the [IP] solution

Test runs have been performed using two real-world test sets of 5 working days each:

1. Set 1: TMS data of 5 working nights (from March 2018, after the dropout of the large consigner) that have not been included in the input data
2. Set 2: TMS data of 5 working nights (from September 2017) are selected as the "worst-case" week, having the highest number of stops in the input data

The routing tool assumes a starting time of 2:45pm and a mean stopping time of $\overline{s^f} = 2.5$ minutes per stop. It computes the schedule and the expected time of arrival per stop. From this information, other statistics can be computed. Fig. 3.10 shows the routing results per district and day in hours for the first test set.

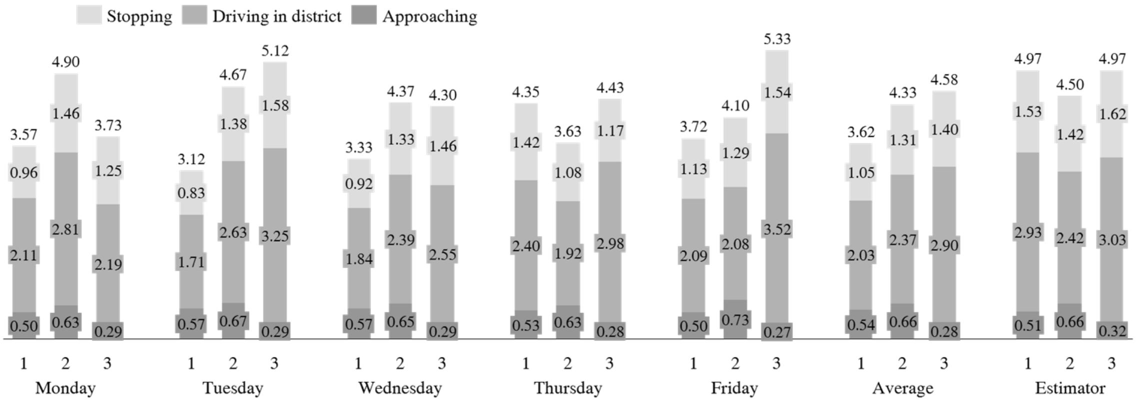


Fig. 3.10 Expected tour duration [h] per district and night in test set No. 1

Comparing the average of the five test days with the estimated values (stacks on the right), one finds that tour 1 has some deviations regarding the driving and stopping time. In the test data, the mean number of stops in tour 1 is 37. Herein, tour 1 has 25.2 stops on average, thus having much less stopping and driving time. Tour 2 and tour 3, however, are within a reasonable range around the estimator. The average speed in the distribution phase from the routing tool was 54.12 km/h for tour 1, 67.82 km/h for tour 2, and 66.62 km/h for tour 3. This confirms that the assumption of $v = 60$ km/h is reasonable.

The results are evaluated to be of good quality, as only one out of 451 stops in the sample was arrived late (delivery after 8am on Friday on tour 3).

The second test set performs much worse as it represents the "worst-case" week, having the most stops on Tuesday. The routing results are shown in Fig. 3.11.

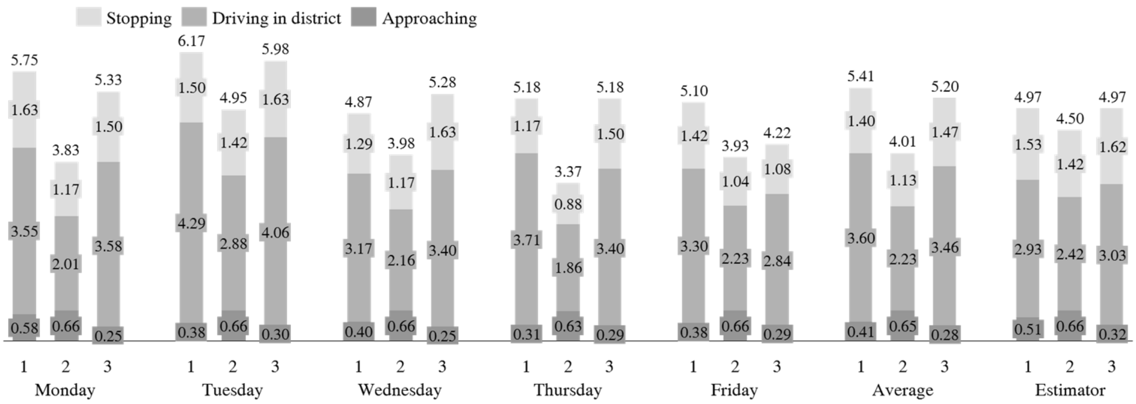


Fig. 3.11 Expected tour duration [h] per district and night in test set No. 2

As one can observe, the results from this "worst-case" week are opposing the results from test set 1. Tour 1 is a heavily utilized tour. In fact, it is over-utilized as the maximal available time is exceeded. Tour 2 is the only tour that has no delayed shipments. Tour 1 exceeds the maximum of 5.25 hours twice, tour 3 thrice. Whereas these delays are trouble, remember that this is the worst case and as such acceptable. As aforementioned in the introductory section: even good districts may sometimes perform poorly due to stochasticity.

Probable reasons for differences in the number of stops are seasonality and other minor changes in the customer structure. For example, spare parts for agricultural equipment are affected by strong seasonality. However, practitioners assured that both test sets are high season periods.

3.5.4.2 Testing the adapted savings and the practitioner's approach

The two alternative solutions from subsection Fig. 3.12 are tested using the test sets 1 and 2, too. Fig. 12 summarizes the results: the estimator is calculated as before. The average values are calculated over the five workdays per test set.

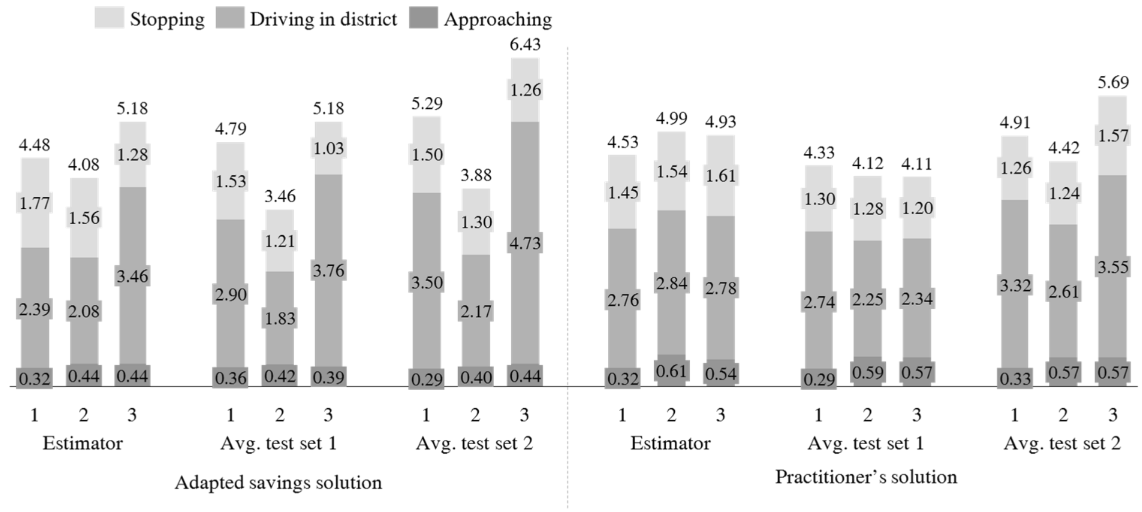


Fig. 3.12 Average and estimated tour duration [h] per district and night of alternative solutions

Both solutions are estimated to be feasible, as all estimated tour durations are less than 5.25h. The savings solution is not very well balanced, the difference between tour 3 and tour 2 is 21.2%. The practitioner's solution (9.2%) is estimated to be slightly better balanced than the [IP] solution (9.4%). The criterion of balance is used as a constraint, not as an objective, in the [IP] formulation. Therefore, this result is not surprising.

The adapted savings solution does not perform well. The average results are less balanced than the estimated durations. In test set 2, tour 3 finishes on average 88 min. too late. The problem stems mostly from an outlier: once it finishes at 11:17am. The tours in this solution have very different tour distances: while tour 2 has an average distance of 173.6km, tour 3 has 355.8km. Nevertheless, tour 3 has a remarkable average travel speed of 67.4km/h. Overall, the savings solution is not compact, it is less balanced, and it has more delayed deliveries than the [IP] solution. Therefore, this approach cannot be recommended. The practitioner's solution performs better: in test set 1, there are no delays at all and the average tour durations are fairly balanced. The difference between tour 1 and tour 3 is only 5%. In test set 2, the difference between tour 3 and tour 2 is 22.3%, which is still better than the [IP] solution (25.9%). Summarising, in test set 2 the [IP] solution performs best with respect to the fewest delayed deliveries. The objective to optimize compactness results in fewer violations and is thus more robust. Since the estimated durations of the practitioner's solution are almost as good as the [IP] solution's estimations and the test results are ambiguous, the practitioner's solution has a great quality, too. The reason for this may be the additional consideration of motorways, which are not considered in the [IP] formulation in section 3.3.

As a result, the [IP] solution is recommended. The focus on compactness comes at the cost of a lower degree of balance. However, especially in peak season the [IP] solution performs more robust, due to the higher degree of compactness.

3.6 Discussion

This paper investigates the districting problem in distribution from the perspective of an INSP. Although there is rich literature on strategic (physical network structure) and operational decisions (daily routing), there is only scarce literature on the districting decision which bridges the gap in-between strategic and operational decisions (van Oudheusden et al. 1999). The described problem is generalizable for all LSP who deliver shipments from a terminal to many stops in a service area. The application fields are parcel services in rural areas, express, and groupage freight. In a broader sense, districting of sales- or repair people are other applications.

3.6.1 Theoretical implications

The contributions are threefold. First, a model for the DDP is presented, which explicitly formulates compactness, contiguity, and balance. Second, the model is combined with a good estimator of travel time. Third, a case study demonstrates the applicability including a parameterization of the model.

Herein, the DDP is formulated as an [IP]. Since most authors do not model the contiguity condition explicitly (Shirabe 2009; Daskin 2010; Kalcsics 2015), this [IP] formulation adopts the net flow approach, which Shirabe (2009) proposes for the general districting problem. Additionally, his net flow idea is used for a new measure of compactness (minimum net flow) in the objective function. This formulation is generalizable for all districting applications that benefit from compactness. If the decision-maker's focus is on a better balance, it should be modelled as the objective (Daskin 2010). Herein, the balance of the computed districts is good, but it is not perfect due to the tolerance δ^{t+}, δ^{t-}. The tolerance is necessary in order to permit solutions that are more compact. In general, there is a trade-off between balance and compactness in every districting decision: the greater the tolerance for the balance, the more room for optimizing compactness and vice versa.

The most important activity measure in distribution is the tour duration because the available time is a critical bottleneck. A good estimator is adopted from literature, especially from Beardwood et al. (1959) and Schmidthöfer (2004). Both literature and experience are used to parameterize the presented estimator. The values for the parameters $\beta = 0.765$ and $\lambda = 1.15$ are general constants from literature. The travel speed v and the stop-time parameters $\overline{s^f}$ and $\overline{s^v}$ may vary considerably from case to case. Deviating estimates change the tour duration and thus the feasibility, balance, and shape of the districts considerably. Therefore, the decision-maker needs either a process model of the delivery process or valid empirical data. Real-time routing showed values of around $v = 60$ km/h and observations resulted in $\overline{s^f} = 2.5$ minutes in the in-night case. These values may not be applicable to other cases, but may serve as a reliable comparison.

In operations research literature, there are only few real-world applications of distribution districting (e.g. Fleischmann and Paraschis 1988). The presented case demonstrates applicability. The DDP is solved computationally using TMS data from Kiessling-Spedition and online navigation data.

It turns out that the [IP] model produces good solutions. One reason may be that the Beardwood et al. formula makes the assumption of uniform density ρ_i of stops. Clustered stops in communities or industrial parks condense the tour and thus improve productivity. Thus, a real tour is probably shorter than the estimated tour. Further, this estimator is especially suited for compact districts having a round shape (Daganzo 1984).

A generalizable advantage of the [IP] solution is the comparably robust result due to its focus on compactness. Compactness grants better routing flexibility in challenging situations. However, the practitioner's rule-based approach performs well, too. In the design of the rules, it is possible to consider some aspects of experience and visual intuition on the map, e. g. connectivity by motorway. That sort of aspects improves the solution quality in daily routing.

3.6.2 Practical implications

Practitioners can draw two lessons from this case. First, how a distribution districting plan should look like. Second, how to get there.

The productivity of distribution districts depends on the driving speed, the stop-time, the density of stops, and the quality of districts. The driving speed depends on the underlying infrastructure (Kellner et al. 2017). The key drivers of the stop-time are the drivers' familiarity with the recipients' depots and the difficulty of deposition. The density of stops is an aspect of the shipment structure. The quality of districts is manageable through districting. Practitioners may need to align their distribution districts whenever the shipment structure changes. Generally, high quality distribution districts have a roundish shape (are compact), do not have gaps, wholes, or overlaps (are contiguous), and the workloads do not deviate from each other (are balanced). The results discourage finger-like shapes of districts because compactness of the districts fosters robustness. Nevertheless, the rule-based approach demonstrated that other intuitive aspects and local geographic knowledge could also improve the solution.

For the planning process, information about the shipment structure, the underlying infrastructure, and the stop-time is a prerequisite. Therefore, the first step of districting is the data collection and an analysis of the shipment structure and the visualization of the results on a map. The IT landscape in Fig. 3.4 supports this managerial process. It is recommended to model the DDP using the presented [IP] or an adjusted variant. Especially large-scale cases benefit from this approach. The next step is a data-driven parameterization of the model. Extensive testing of the solution using a real-time routing software reveals shortfalls of the model in practical operation before the implementation. At this point, other prevalent arguments (see subsection 3.5.2) can be considered through manual adjustments.

3.6.3 Limitations and further research

The approach taken in this paper has its limitations. First, the model considers uncertainty of varying stop count and location by using the average over several months. This ignores the variance over time. As a result, the model is virtually deterministic. Further research should develop an activity measure of uncertainty to improve the robustness of the solution.

Second, although the parameterization of the model is done meticulously, it may be further augmented. Estimating the right driving speed - a key driver of productivity - of an area drives the accuracy of the travel time estimator. A big data approach like in Kellner et al. (2017) should be adopted to improve the input data. Tests have shown that the average overland speed (61-70km/h) is significantly greater than the urban speed (41-50 km/h).

Third, what is called the practitioner's approach is not yet formalized as a solution procedure. The rules with respect to motorways and the allocation of major towns have produced good results. Herein, both approaches are compared, but not yet integrated. Further research on this integration could translate these rules into a formal heuristic. This is considered to be worthwhile in order to strengthen the bridge between theory and practice.

It has been elaborated that compactness of districts fosters routing flexibility. Managerial approaches that preserve mid-term stability and exploit operational flexibility should be investigated. That sort of operational measures could further improve the quality of districts.

4 Bringing infrastructure into pricing in road freight transportation – A measuring concept based on navigation service data

Authors: Florian Kellner, Andreas Otto, Christian Brabänder

This chapter is the peer-reviewed and published version of Kellner, Florian, Andreas Otto, and Christian Brabänder "Bringing infrastructure into pricing in road freight transportation–A measuring concept based on navigation service data.", in: Transportation Research Procedia, 2017, Vol. 25: pp. 794-805, DOI: 10.1016/j.trpro.2017.05.458.

Abstract

Differences in road infrastructure, such as capacity, congestion, speed limits affect the productivity and the costs of short-distance freight operations. This article introduces a novel methodology that is based on navigation service data to measure the effects in terms of cost per kilometer using navigation and shipment data. The methodology is applied to five terminals of a forwarding cooperation and has been able to document significant differences in cost per kilometer across the terminals.

Keywords

Traffic congestion; Road freight transportation; Logistics service provider; Navigation service; Short-distance freight transportation

4.1 The need to measure and quantify infrastructural differences in transport operations

Road infrastructure, all assets of the navigable road network (roads, traffic lights…) and the experienced traffic conditions, determines the productivity of freight transport operations, which can be measured by stops per day of a vehicle (Figliozzi 2010). If a vehicle does 13 instead of 14 stops due to poor infrastructure, the productivity decreases by 7%. Most metropolitan areas are prone to congestion. A provider of GPS assisted navigation systems, for example, ranks Stuttgart, Hamburg, Berlin, and Munich (in this order) top on a traffic index with the highest traffic problems across Germany (TomTom 2014). Freight operations, if measured in cost per kilometer, in these cities are more expensive than in rural areas. The question is how much more expensive? If the effect cannot be quantified, cost accounting and pricing cannot process it – although it may promote better decision-making.

This article addresses two research questions: How can infrastructural differences get measured? How can they be translated into cost differences?

The ability to measure these effects is important in several situations: (1) Pricing: It promotes transparent and regionally differentiated pricing of logistics services. Regions or single customers may be more expensive to serve. (2) Fair cost allocations in a cooperation: Less-than-truckload forwarders run networks. One terminal collects shipments, another delivers shipments in another area. The process is usually covered by internal transfer pricing, which reimburses the delivering terminal. The reimbursement should reflect infrastructural differences since they affect productivity. Correct cost allocations are a goal in itself. Furthermore, if ignored, the stability of a less-than-truckload cooperation is at risk, since the delivery operation either yields too much or too few profits, increasing the propensity to leave the cooperation. (3) Location decisions: Real estate companies may use the method to evaluate the relative advantage of assets. Location decisions, for example for distribution centers, can incorporate it into their decision support model.

This paper introduces a methodology to measure the effects of infrastructural differences in terms of cost per kilometer using data from navigation service providers (e. g. Bing Maps, Google Maps, INRIX, HERE, TomTom …). In a computational application the methodology is applied to five terminals of a forwarding cooperation. The paper closes by listing limitations and issues for further research.

Navigation and traffic information providers dispose of a detailed mapping of road networks and combine different techniques in order to offer real-time information about the current traffic situation throughout the network. As the average travel speeds on the different network segments may vary significantly during the day depending on the local traffic conditions, it is important to consider the time of day when the transport operations are carried out. For that reason, 'static' time-matrices that do not take time-varying travel times into account are not used in this research when calculating the time-based cost component of freight delivery.

4.2 Literature review

Several aspects of measuring infrastructure differences and translations into cost-effects have been researched in literature and are relevant to this paper.

Addressee: Research in the fields of economy and geography views congestion costs as a loss to public welfare. Investments in projects attempting to mitigate congestion should be weighed against this loss (Public Policy Institute of California (PPIC) 2002). Schrank and Lomax (2002) estimated the costs of lost time and fuel caused by congestion at 67.5 billion USD in the USA for the year 2000. Traffic congestion also causes people to have higher stress levels ("Commuter Stress Index"; Wener et al. 2005) and has effects on health (Bigazzi and Figliozzi 2012, 2013; Levy et al. 2010). All studies expect increasing costs caused by congestion (Sarzynski et al. 2006). Most research considers the effects of congestion on the public. This paper analyzes the effects of infrastructure on transporters' costs. Eisele and Schrank (2010), Figliozzi (2010), or Taylor et al. (2012) took the same perspective, but for different purposes than pricing.

Data source and data supplier: Information on traffic congestion can be measured using stationary sensors, induction loops, image recognition, signals from moving vehicles ("Probe Drivers") or can be estimated with indicators such as traffic lights, etc. Lomax et al. (Lomax et al. 2011) use the data of a navigation services provider. Travel times, offered by navigation service providers exploit different techniques to observe local traffic conditions, including induction loops and signals from moving vehicles. Time-varying travel time information can be accessed in real-time. The authors regard this as a fundamental improvement ("Game Changer"; Lomax et al. 2011). This approach is adopted in this paper.

Indicators to measure congestion: Rao and Rao (Rao and Rao 2012) review the use of the indicators speed, travel time / travel time delay, vehicle volume, "level of service" (quotient of the volume and capacity of the vehicle), and increase of costs. Lomax et al. (2011) determine a Travel Time Index (TTI; max. travel time / travel time in fluent traffic), also described as "speed reduction factor," as a key indicator, but also suggest other indicators (travel speed, travel delay, annual person delay, annual delay per auto commuter, annual peak period travel time, travel time index, commuter stress index, wasted fuel, total congestion cost and truck congestion cost, truck commodity value, roadway congestion index, number of rush hours, percentage of daily and peak travel in congested conditions, percentage of congested travel). The authors also analyze freight transportation and determine "truck congestion cost" as the product of time spent in congestion and costs per hour (Lomax et al. 2011).

Goals: The majority of existing research elaborates on congestion within specific areas. Several contributions connect the results with recommendations. Most of them suggest applying toll and a need for further infrastructure investments. An overview is presented by Parry (Parry 2008).

The research questions raise the necessity to define a comparable indicator that makes different regions comparable in terms of cost-effects. Literature considers several approaches to measure traffic congestion in a macroeconomic perspective. This research reveals that general congestion indices, which are not based on specific companies, can lead to wrong evaluations. For example, Stuttgart does not necessarily have to be on top in a German-wide rating, although industry experts would intuitively expect a top rating. We are not aware of existing approaches to exhaust navigation data in cost measuring for freight transportation with the purpose of pricing. This essential shortcoming from a network perspective will be tackled in the course of this essay. We believe that there are fundamental differences in the infrastructure, where transporters operate. As a result, one needs to account for infrastructure in the pricing of freight transportation. In order to develop such an indicator we propose to follow a "Big Data" approach. Navigation service data is used by thousands of individual vehicles in every region on a daily basis. Using this data is a straight-forward approach to compare infrastructures. One has to develop a smart way of exploring this data and transform it into cost factors. We are presenting a data exploration concept that first extracts specific instances from navigation service data and second derives cost rates, which are comparable across regions.

4.3 Conceptual foundations of the method – requirements for measuring

4.3.1 Static and dynamic attributes

Infrastructure includes all facilities and factors, which maintain the basic functions of an economy (Frerich and Müller 2004). A part of it is road infrastructure, consisting of assets (roads, traffic lights, …) and rules (driving restrictions, parking restrictions,…). Infrastructure has static and dynamic attributes. Static attributes are for example the length of the road network, the width of streets, speed limits, or driving restrictions. Dynamic attributes change over time, where congestion is a dynamic attribute which is central to our reasoning. As both factors affect the productivity this paper suggests processing static and dynamic attributes.

4.3.2 General indices are not applicable

Several sources report congestion indices for cities. These cannot be used for the problem at hand. (1) Coverage: First they report per city, where a report is needed for particular service regions, consisting of cities and rural areas. (2) Relevance: A congestion index may report a high value for a particular city within the service region. However, if there are no/few locations to be serviced, the congestion is irrelevant.

4.3.3 Measuring based on "weighted costs per kilometer of shortest distance"

It is suggested to use "weighted costs per kilometer of shortest distance" as an indicator to quantify differences in road infrastructure. It is relevant to distinguish between the shortest distance, the quickest distance, and actually travelled distance, since congestion might cause the driver to take detours; hence, costs per kilometer decrease due to higher speeds during the detours. Given the shortest distance, differences in infrastructure cause both varying driving times and a varying number of covered kilometers. Thus, costs per kilometer of shortest distance will vary proportionally.

4.4 Collecting navigation service data

It is suggested to identify differences in infrastructure by using data from navigation service providers. Put simply: average speed for defined segments is measured by retrieving data from a provider of navigation services. The costs are then calculated as a function of distance (km) and travel time as obtained from the provider. The following elaborates on this.

4.4.1 Determine what to measure

It is necessary to define which data to collect. In principle, three options are available. These will be evaluated with respect to three criteria: representativeness, data supply availability, and data processing capabilities.

4.4.1.1 "Real World" tours

Delivery and pick-up is done in tours where a tour is defined as the number of shipments that are delivered or picked up with one vehicle in one journey and a route is defined as the sequence of the locations serviced along during the tour. A tour with the actual route can be simulated by a navigation service provider reporting trip length, travel time, and delay. Average speeds can be calculated, low average speed indicates poor infrastructure. This option is representative since it includes all stops and distances travelled during a certain time period. In real-world tours the routes are given. It is assumed that the driver optimizes the route based on experience and local knowledge, avoiding congestion.

Nevertheless, this option is not selected since the data collection requirements are considered prohibitively high. A less-than-truckload network in Germany, for example, typically has 40 to 45 terminals. Each terminal would need to report the data of the real-world tours. Furthermore, each tour would have to be simulated by the navigation service. 40 terminals, each having 50 tours per day, would lead to 2,000 tours just for a one day report. However, since many customers will only occasionally receive shipments, a one day report is not representative and would have to be repeated for several days. Thus, this option was not selected.

4.4.1.2 Artificial tours

The second option is to create "artificial" tours based on actual location and shipment data using a tour planning algorithm. This option is advantageous, since the needed shipment and fleet data is considered easier to be retrieved. But the requirements for data processing are prohibitively high, since the tours have to be build. Past experience shows that calculations, even on moderately long tours with 12 to 15 stops, are very time-consuming. Again, for a single day, 2,000 tours would have to be build. Furthermore, artificial tours will be less representative, since the tours and routes may be unrealistic.

4.4.1.3 Tour segments

This option neither uses real nor artificial tours, but uses rules to create tour segments, where a segment connects two customer locations. So, instead of simulating and measuring the transport operation as closely as possibly by real-world or artificial tours (which has been rejected as explained above), we suggest building a representative set of segments. These are then evaluated by the navigation service. The costs associated with the quickest path for a given segment are divided by the distance on the shortest path. For a given day, only a limited number of customers are selected to build the segments. High volume customers have a higher probability of becoming selected.

This option is prima facie less representative since it samples, it creates segments that may never be driven, and may exclude congested segments, which cannot be avoided.[43] However, taking into account that both real-world and artificial tours are only one-day (or few day) snapshots of an "all-year" reality, the suggested volume-based sampling of segments can yield a better representation. Furthermore, the data requirements are as low as for the "artificial tours" and the processing requirements are considerably lower since tour building is not needed. Thus, it is suggested to continue with "tour segments".

4.4.1.4 Two types of segments: "Approach" and "Delivery"

The measuring concept differentiates two types of segments: approach and delivery. The "approach" represents the segment that connects the terminal and the delivery area (represented by an "anchor", see below). The "delivery" represents the segments connecting customers within the delivery area.

Segment type "approach": The "approach" is the quickest connection between the terminal and an anchor. The anchors are determined like this: A terminal serves an area ("terminal area"; typically defined by zip codes). A terminal area consists of *n* delivery zones, where each delivery zone has one anchor. An anchor is also a customer. The parameter *n* should resemble the number of tours actually driven on average in the terminal area. The determination of the anchors and the assignment of customers to delivery zones (Multi-Weber problem (Plastria 2011; ReVelle and Eiselt 2005)) are done using a k-means clustering algorithm (Cooper 1964), where the shipment volumes per customer are used as weights. The algorithm results in a pair of continuous coordinates. The customer nearest to this coordinates is the "anchor". Distance and travel time between terminal and anchor are measured every *min* minutes.

[43] However, the mistake of excluding segments "to" and "from" the terminal cannot occur (see below).

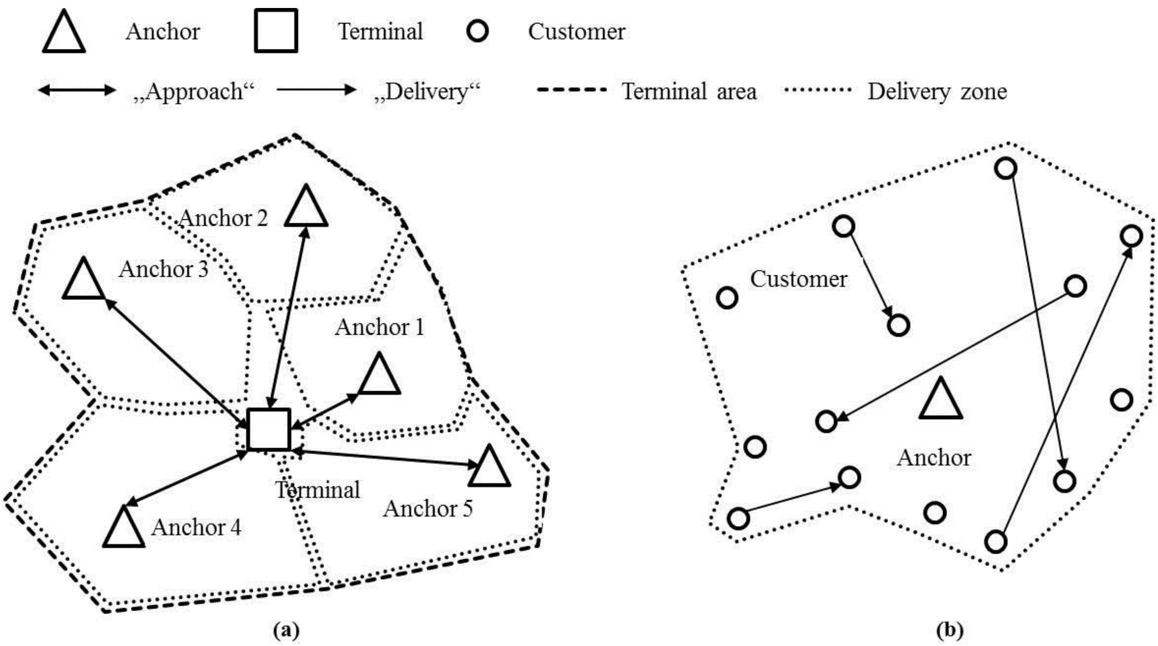


Fig. 4.1 (a) Segment type "approach"; (b) Segment type "delivery".

Segment type "delivery": Within each delivery zone, *k* segments (customer to customer connections) are created. The selection of customers is carried out randomly, excluding customers with less than one delivery per week. Distance and travel time are measured every *min* minutes. The parameter *k* is the result of some simulations. The variability of the results increases for smaller *k*. Repeated random selection and calculation for parameter $k=20$ leads to a good trade-off between reliable results and computational demands.

Weighing of the segment types: Each terminal is represented by two series of measurement (approach and delivery). These were merged using weighting. It is suggested to weigh according to the share of the total driving time per day across both approach and delivery. A delivery zone, which is far away from the terminal, has a long approach. If the costs per km are high, the influence is stronger, if the approach has a long distance. If the approach is short, the delivery in the delivery zone gains higher weight. Therefore, if a vehicle travels a long way from the terminal to the delivery zone *a*, the relative weight of the approach $g_{terminal,approach}$ is high, if it is a short way, $g_{terminal,approach}$ is smaller. The duration of the approach (and the return) is estimated for the delivery zone *a* with the requested travel time from the terminal to the anchor of *a*. The delivery time is the daily travel time minus the time for the approach and return. The segment types are included in the final indicator $\delta_{terminal}$ based on those times.

$$\delta_{\text{terminal}} = g_{\text{terminal,approach}} \cdot \left(\frac{C_{\text{terminal,approach}}}{D_{\text{terminal,approach}}} \right) + g_{\text{terminal,delivery}} \cdot \left(\frac{C_{\text{terminal,delivery}}}{D_{\text{terminal,delivery}}} \right) \quad (1)$$

with:

δ_{terminal}	weighted average cost rate for the terminal area
$g_{\text{terminal,approach}}$	relative weight of "approaches" segments in the terminal area
$g_{\text{terminal,delivery}}$	relative weight of "delivery" segments in the terminal area

$C_{\text{terminal,approach}}$ sum of the costs of all "approach" segments' travel time and distance in all delivery zones *a* (on the quickest route) over all points in time observed

$C_{\text{terminal,delivery}}$ sum of the costs of all "delivery" segments' travel time and distance in all delivery zones *a* (on the quickest route) over all points in time observed

$D_{\text{terminal,approach}}$ sum of the kilometers traveled on all "approach" segments into all delivery zones *a* (on the shortest route) and vice versa over all points in time observed

$D_{\text{terminal,delivery}}$ sum of the kilometers traveled on all "delivery" segments in all delivery zones *a* (on the shortest route) over all points in time observed

The weights are calculated according to equations (2) and (3):

$$g_{\text{terminal,approach}} = \frac{t_{\text{terminal,approach}}}{t_{\text{day}}} \tag{2}$$

$$g_{\text{terminal,delivery}} = \frac{\left(t_{\text{day}} - t_{\text{terminal,approach}}\right)}{t_{\text{day}}} \tag{3}$$

with:

t_{day} pure daily travel time

$t_{\text{terminal,approach}}$ average travel time of "approach" segments plus return in the terminal area over all delivery zones *a* and points in time

4.4.2 Necessary navigation data

A request is defined as the technical process of sending a query to an application programming interface (API). The query contains starting coordinates (from) and destination coordinates (to) and specifies the request in terms of whether the shortest or the quickest path is requested, and whether travel times should represent historic data (e. g., average travel time on Monday, 5 p.m.) or current data (traffic situation at the moment of the request). Historic data basically promotes representativeness. However, the use of averaged observations is not useful if these observations refer to a whole day or several days, as irrelevant periods are processed.[44]

In real life applications, the number of needed requests will typically be very large, in order to produce reliable results. For the "approach" each terminal *m* records outbound and return trip information to and from n anchors at t points in time per day. For the "delivery", data for k segments is collected for each delivery zone *a* at t points in time. The total number of requests per day nr is:

$$nr = t \cdot m \cdot n \cdot 2 + t \cdot m \cdot n \cdot k = t \cdot m \cdot n \cdot \left(2 + k\right) \tag{4}$$

with:

nr total number of requests per day

[44] Consider the situation where the distance between the terminal and a certain delivery area is about 80 km. Typically, that delivery area will not be reached before 8 a.m. and it will be left before 15 p.m. Any congestion outside that period should not be integrated in the measurement.

t	number of observations (points in time) per day
m	number of terminal areas
n	number of anchors per terminal area
k	number of customer-customer segments per delivery zone

The first summand in equation (4) captures the number of requests for the segments connecting the terminal with the anchors; the second summand captures the number of requests for the delivery segments.

4.4.2.1 Timing of measurements

The approach and segments within the delivery set should not be measured throughout the day. Congestion far away from the terminal in the early morning will be irrelevant since this location will only be reached around noon. Thus, the routine that selects customers of an delivery zone to be part of the delivery set, has to rank the customers according to the time of the day.[45]

4.4.2.2 Frequency of measurements

When travel time per segment varies seasonally, the frequency of measurements becomes important. Ideally, distances and travel times are recorded on each weekday throughout the year. Yet, data handling will become difficult. The frequency of measurements should be determined as a function of the (potentially) identified variations of travel times. It is especially relevant to represent the overall variations of travel times in the infrastructure completely and, at the same time, to ensure that the key indicator stabilizes.

4.5 Application: Comparing five German regions in a freight transportation network

The concept has been applied to an existing German-based less-than-truckload freight-forwarding alliance, here abbreviated "FA". FA's goal has been to report infrastructural-related differences in terms of cost per kilometer for each terminal. This data is used to amend transfer pricing between terminals. The current pricing is based on a price matrix across all terminals driven by shipment weight and distance (terminal to customer). It should be amended by a "cost per kilometer"-index per terminal. FA considers this helpful, since the terminals in high-congestion areas are not properly compensated. FA-members reported that a terminal's ability to generate profits depends on infrastructural attributes.

[45] More precisely: The most likely timeframe for the vehicle to be present at a certain segment has to be determined. This timeframe will depend on numerous conditions (exact routing, time windows, …), amongst others on the distance between the customer location and the terminal. The probable timeframe when the vehicle is present at a certain customer should be determined as a function of the distance between terminal and customer location as this is the only information that is generally available for each customer and delivery area.

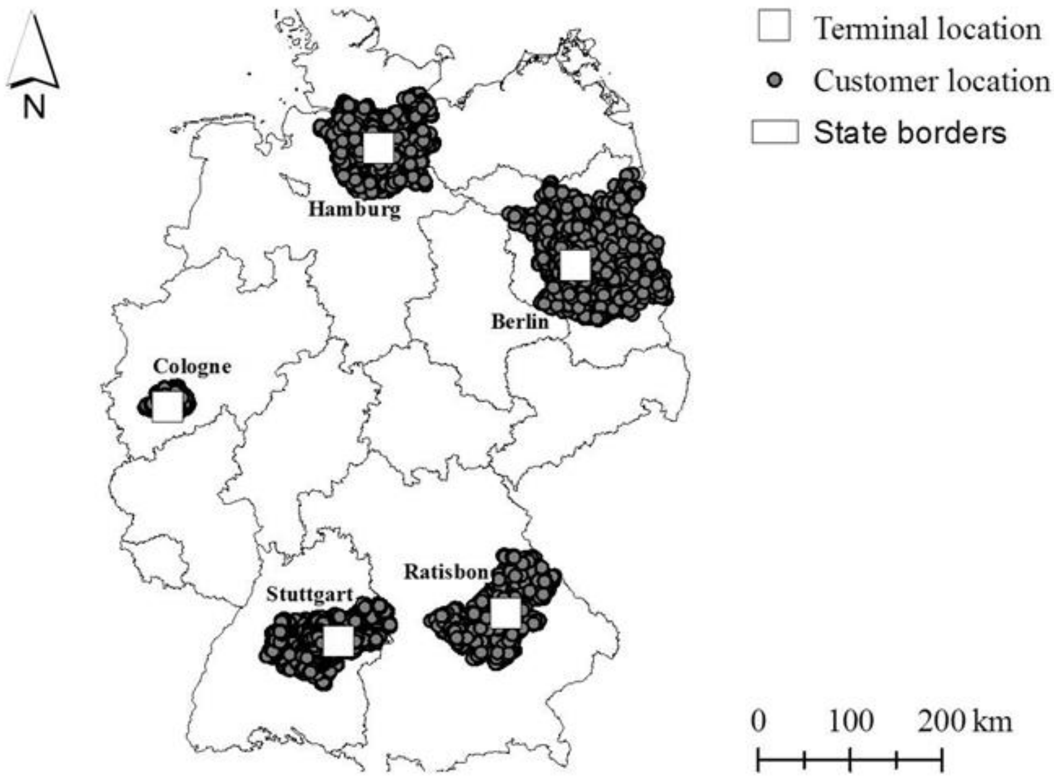


Fig. 4.2 Terminals and terminal areas

FA has 45 partners in Germany, i. e. runs 45 terminals. The terminals Berlin, Hamburg, Cologne, Stuttgart, and Ratisbon have been evaluated. Data on terminal, customer location, and shipment volumes have been obtained. Fig. 4.2 displays customer locations (dark bullets) across Germany. In total, 50.105 unique customer locations have been included in the measurement. The white squares represent the terminal locations.

For each terminal the k-media clustering algorithm has been applied. Fig. 4.3 displays the Ratisbon terminal (white square) and the anchors (white triangles).[46] Per terminal area and delivery zone $k =$ 20 segments have been selected.

4.5.1 Measurements

Data for "approach" and "delivery" have been collected from 6 a.m. to 9 p.m. on two consecutive days with an interval of 20 minutes. So, 15 hours x 3 observations/hour = 45 observations have been done.[47] The quickest travel time and the respective distance were determined for each segment. The shortest distances were collected with a separate query. Since thousands of segments had to be measured, the query applications took a few minutes (about 7 minutes for all queries).

[46] One observes a strong concentration of anchors in the north of Ratisbon due to few but very frequently approached customers.
[47] Measurements on additional days will increase the reliability of the results. However, since navigation service providers charge per request, this application has been restricted to a request volume that allows verifying the concept.

The data points were interpreted as simultaneous, i.e. denoted by the starting point of each interval. As expected, the shortest distances between two points have been constant, while the quickest travel path varied both in travel time and travel distance across the measurements.

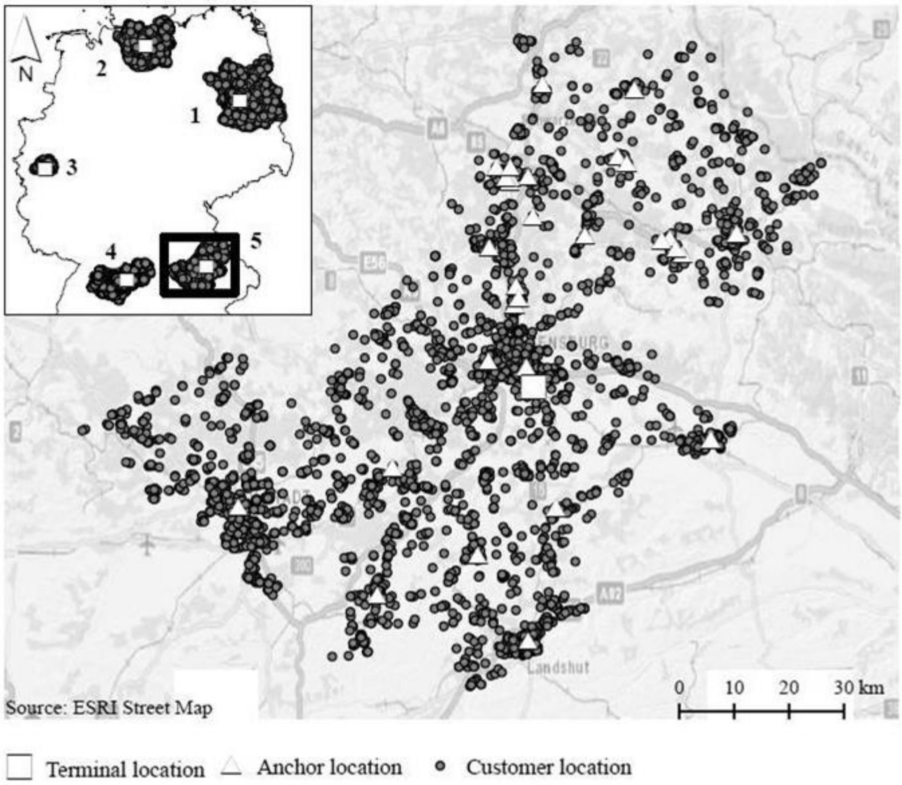


Fig. 4.3 Terminal Ratisbon (square) with customers (bubbles) and anchors (triangles)

4.5.2 Computation

Segment type "approach": For each delivery zone, travel times and traveled kilometers have been multiplied with cost factors per km and per hour. The total sum of the costs was divided by the sum of shortest distances. As cost data, 24.11€ per hour of vehicle usage and 0.50€ per km driven has been applied. These are in line with literature (Wittenbrink 2011).

Segment type "delivery": To compute the costs per kilometer of shortest distance of the "delivery", the selected segments within the clusters were valued with the costs per km and hour. The sum of costs per terminal was divided by the sum of "shortest" distances on the same segments. Again, this is necessary to ensure consistency.

4.5.3 Computational results

The estimated cost rates per km of shortest distance range from 1.06€ (Ratisbon) to 1.38€ (Cologne; Fig. 4.4 a). The relation between "approach" and "delivery" depends both on the exact location of the terminal and the customer locations and the overall size of the area. For example, the smallest area (Cologne) has short distances to approach the anchors and delivery tours that remain in urban areas (no highways, no rural areas). On the other side, terminal Stuttgart is located beyond the inner urban area. Therefore, approaching most anchors is not affected by urban traffic. Thus, the computed costs of 1.14€ per km are low compared to pure metropolitan locations, like Cologne. Fig. 4.4 (b) shows the costs per km relative to the mean across the terminals.

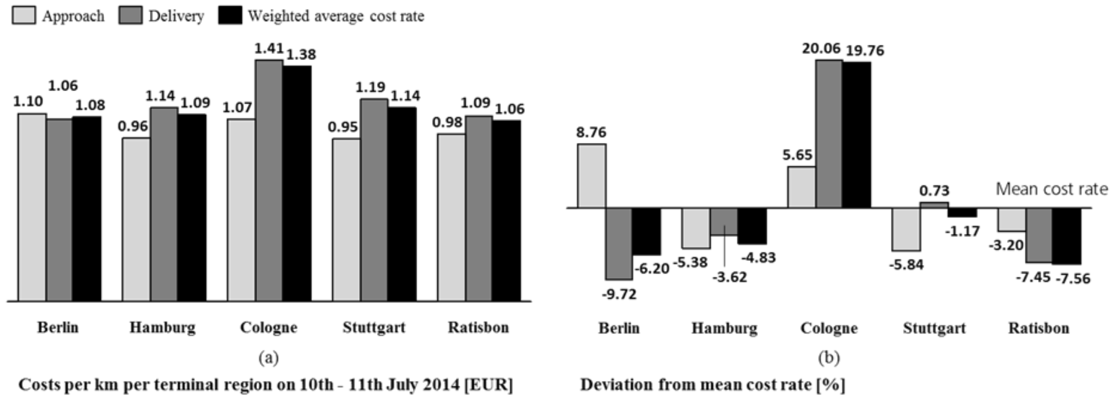


Fig. 4.4 (a) Costs per km of shortest distance in [EUR], as shown in formula (1); (b) deviation of costs from mean of costs per km across all terminal areas [%].

The results suggest that Cologne, for example, should receive an increased transfer price due to less favorable infrastructure where Ratisbon would have to accept the largest reductions.

4.6 Limitations and further research

4.6.1 Internal and empirical validity

Internal validity is an attribute of an observation. It asks whether what has been measured has been intended to measure (Frankfort-Nachmias and Nachmias 2008). In this case, the hypothesized relationship between dependent costs and independent variable (infrastructure) is to be measured. Arguing strictly, navigation data does not report infrastructure differences but only their effects. For example: A road with a traffic light reduces average speed. The navigation data does not report the traffic light (as an element of infrastructure) but only the effects (speed). This may be considered a weak link in the proposed concept.

Empirical validity is an attribute of an observation. It asks whether it is possible to draw conclusions that match reality. The proposed concept needs empirical validation, which may be pursued by evaluating driving reports as produced by telematics devices in the vehicles. These allow verifying driving speed and stop times for particular routes.

4.6.2 External validity

External validity is an attribute of an observation. It asks whether the observation represents the population. The proposed concept observes per terminal area only selected segments at selected times, thus is only a fraction of reality. It needs to be studied in more detail how variations in number and selection of segments affect the results.

Furthermore, external validity may be curbed by an insufficient number of observations. The current results are confined to only two working days. Fig. 4.5 reports on the relative variation [%] from the 2-day-mean. The spread of regional cost rates remained inside a narrow interval [-1.02%; 1.00%] around the mean. In other words: The 1-day indicator does not significantly deviate from the 2-day indicator.

Eventually, there have been deviations of 1 Eurocent for the terminals in Stuttgart and Hamburg between day 1 and 2. For the other three terminals Berlin, Cologne, and Ratisbon the cost rates on both days have been identical. In future applications it will be necessary to repeat the study multiple times over a long run, in order to establish conclusions about reliability of the concept.

Fig. 4.5 2-day comparison of reliability [%] in the application case.

4.6.3 Efficiency

The concept requires a lot of data requests. These data requests are the major cost driver of the proposed method. Therefore, further research should study how to reduce the number of data requests necessary. There are several starting points.

One way would be to cooperate with navigation service providers. A single request should generate a response containing multiple historical travel times (e. g. request of all travel times across a past day on one path) and distances at multiple points in time. Within the same agreement, one request could refer to a complete distance matrix of real-time data instead of one path (e. g. define a distance matrix of 1,000 relevant paths' in Ratisbon and request real-time data for this matrix with one single request). Second, further research has to be conducted on options to decrease the total number of necessary data instances. In other words: Which data instances can be neglected without loss in quality? Third, another research question will be how to aggregate customer locations and thus reducing the required number of representative segments to be measured (e. g. aggregation of a neighborhood of ten customers may reduce the number of total locations and thus necessary requests)

5 Final Conclusion

DSS integrate data and support its organization and analysis for an economic purpose. Based on data analysis, DSS find, outline, and evaluate decision-alternatives. The more accurate and reliable the data is, the better are the model's capabilities. Hence, this dissertation approaches the cost issue of the last mile using an analytical, model-based decision-support approach. The basic idea of that approach is to model the last mile transportation process as a graph and determine the values of decision variables that optimize some quantitative objective. The paradigm of the presented models is the flow of CG from a BBP through that graph to many points of delivery. The practical purpose of the models is the efficient generation of transportation services on the last mile. Efficiency means leveraging economies of transportation on the last mile.

5.1 Implications

The practical and theoretical implications are elaborated separately in the three papers. The following sub-section summarizes those implications and connects the dots to a broader picture of last mile transportation.

5.1.1 Implications of focus 1

Focus 1 integrates the vehicle routing decision, the selection of common carriers, and the clustering ("picking") decision of self-fulfillment and outsourcing. This integration of decisions contributes to challenge 1 of Fink et al. (2014). In daily operation, LSPs are contracted to distribute a great number of shipments. For example, a manufacturer of FMCG contracts a large LSP to dispose the production. Therefore, the contracted LSP will collect the bulk of shipments and feed it into its FTL distribution network. In the regional terminals (BBP) of the network, the LSP breaks up this bulk and sorts the shipments, selects the shipments for self-fulfillment and the shipments for subcontracting. Typically, the small, far-off shipments are outsourced to CEP providers. The self-fulfillment shipments are then routed to their destinations in the service area. The fleet of an LSP has a limited capacity. Due to this constraint, LSPs acquire (rent) additional capacities in the short-term to adapt flexibly.

The first paper captured this typical situation with an integrated linearized MIP model. It further shows that VNS-based heuristics are capable to solve that sort of integrated problems in short computation time. The summarized implications from this paper are:

- As expected, the own capacitated fleet is always used first. This is a result of the input factors: the fixed cost of a private fleet are no decision variable as the strategic fleet sizing problem is not considered here.
- The private fleet has been downsized (sub-section 2.5.1, Table 2.2) to test the economic benefits of the flexible rental and subcontracting options. A comparison of the VRPPC and the VRPPCdR test instance from (Bolduc et al. 2007) shows a relative cost reduction of 13.6% due to the newly available options. However, the largest cost reduction for a specific scenario is not only achieved with the scenario including all possible carrier and rental options, but also with scenarios offering only a subset of all options. The new operational

flexibility has great potential to reduce costs significantly, but does not guarantee it. As a result, a large range of subcontracting options does not reduce costs superiorly, but the right subset of options does. Flexibility does not create value per se, it needs careful management.

- The carriers offering volume discounts (carriers C3 and C4, Fig. 2.2, Table 2.6) are favored in the carrier selection but they receive mostly singular shipments in the outskirt of the service area. The effect of volume discounts is expected to be a matter of dispersion. The more dispersed the shipments are, the more shipments will be subcontracted. Further research should integrate the selection of subcontractors with the fleet sizing to establish a reasonable capacity of the private fleet.

5.1.2 Implications of focus 2

The focus 2 of this dissertation takes on the distribution districting problem. The model-based districting decision contributes to the second challenge of Fink et al. (2014). The model focuses on consistency and on robustness in order to deal with the uncertain number and location of stops. The most important, but not the only, application of this paper is the last mile of CEP services. As such, the case of Kiessling's in-night express distribution in the central Oberpfalz is presented.

The summarized implications from the second focus are:

- Districting as a general problem in service science and distribution districting is an ever-repeating problem on the last mile. LSPs need to evaluate and align their distribution districts whenever the shipment structure or the fleet changes. Generally, high quality districts are compact, contiguous, and balanced.
- It is recommended to model the DDP using the presented Integer Program or an adjusted variant. Especially large-scale cases benefit from model-based districting decisions. Adjustments are appropriate with respect to (a) the activity measures and (b) the selection of either compactness or balance as the objective function.
- Compactness has led to superior robustness in the "worst case" test set (Fig. 3.11). This is interpreted as a confirmation that compact districts are a suitable measure for robust last mile distribution. The results discourage finger-like shapes and discontinuous districts.
- The trueness of the districting model depends strongly on the parameterization. The IT landscape in Fig. 3.4 supports this pre-processing. Decision-makers demand information about the shipment structure, the underlying infrastructure, and the stop-time.
- The post-processing of the model's solution includes visualization on a map, testing of the solution using a vehicle routing engine, and feedback from operative employees. The rule-based practitioner's approach demonstrated that intuitive aspects and local geographic knowledge can also improve the solution. As a result, it is encouraged to adapt model-based districts in the light of long-term local experience and prevalent concerns (see subsection 3.5.2).

The DDP has a tactical planning horizon in order to achieve more consistency. The consideration of consistency sacrifices operational flexibility of daily optimized tour clustering. Vehicle routing is performed on a daily basis inside the districts but not across district boarders. The implications from focus 1 are nevertheless useful because the operation of districts benefits from optimized daily vehicle routing and subcontracting decisions.

5.1.3 Implications of focus 3

Focus 3 improves the quality of input data representing the cost of transportation. The methodology proposes to use the "weighted costs per kilometer of shortest distance in [EUR/km]" as the key unit of measurement (sub-section 4.3.3). Nevertheless, the methodology is capable of measuring other units, such as "average speed on fastest way in [km/h]" or "driving time per km on shortest distance in [sec]".

The implications of the third paper are:

- Infrastructural differences should be measured based on real-time navigation data in order to isolate the effect of the underlying infrastructure from the effects of different assets having different productivity and from different shipment structures.
- The weighted cost per km of the shortest distance varies considerably. However, from the five investigated terminal areas, the Cologne area seems to be an outlier. The variance of the other four areas is significant but less than 10%.
- The number of requests for data is very large. A meaningful selection of the right time-windows and the right edges to measure is mandatory to ensure the accuracy of the estimation.

5.2 Further research

Starting points for further research are elaborated separately in the three papers. The following sub-sections propose further research that arises from the integration of the three focus papers. In addition, the last sub-section 5.2.4 returns to the very first sub-section 1.1 and points out other promising approaches to the issues of the last mile.

5.2.1 Integration of routing, districting, and subcontracting

The integration of vehicle routing and subcontracting districts and shipments needs further research. An LSP who operates an own fleet in districts often subcontracts some districts in practice. This means that there is a division of the service area into a self-fulfilled service area and a subcontracted area. The subcontracted area can be districted even further into multiple subcontractors' districts. This districting problem needs integration with the routing problem in order to leverage economies of transportation in the self-fulfilled area. The basic idea is visualized schematically in Fig. 5.1. While the VRPPCdR integrates routing and subcontracting on an operative planning horizon, this proposal integrates an operative and a tactical problem. Wong and Beasley (1984) demonstrate how to make a virtue of this necessity. They solve an operative routing problem and identify the most frequent commonalities.

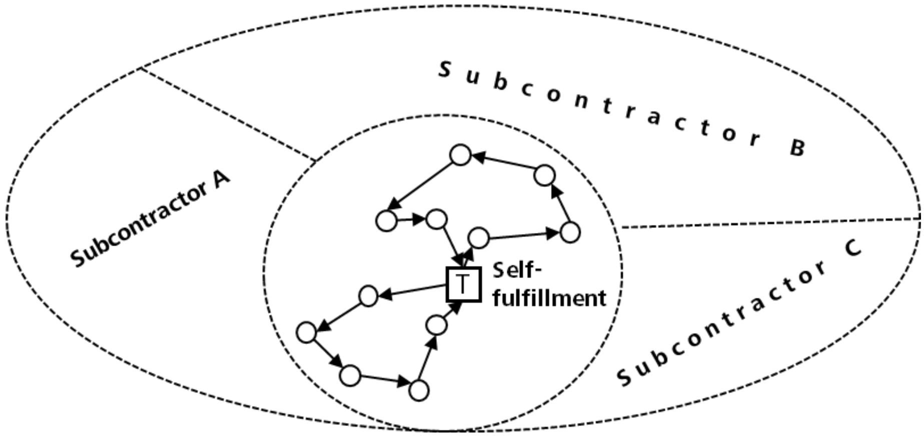


Fig. 5.1 Schematic integration of routing own vehicles and subcontracting districts

5.2.2 Understanding the perspective of the subcontractor

In focus 1, the subcontracting decision is integrated with the routing decision. This paper took the perspective of an LSP (or self-fulfilling shipper) who subcontracts common carriers. The typical case is the subcontracting of CEP providers for the small, far-off shipments having low spatial density. The perspective of the common carrier offering volume discount needs further research. The basic idea of offering volume discounts is to foster outsourcing of the “better” shipments that have higher density and higher volume. Since that sort of “better” shipments enables better densification and thus productivity, the common carrier can offer better discounts. The pricing, including discounts from the perspective of a subcontractor, is a topic for further research. Further, the pricing decision benefits from a sophisticated valuation of the outsourced shipments. First, the value of a shipment differs from structural differences of the subcontractor’s shipment structure: “Does the shipment blend in smoothly?” Second, the value of shipments differ because of infrastructural differences: “How much is it to approach this destination?” This could be measured using a methodology like in focus 3. Districting can support that pricing decision. A common carrier operating districts can average the value of shipments within a district.

5.2.3 Bringing infrastructure into routing and districting

In focus 3, the authors have an allocation problem in mind (sub-section 4.1). The costs of last mile distribution are allocated to the outbound and the inbound members of a GF alliance using transfer prices. The transfer prices are corrected for the infrastructural differences. The presented methodology is applicable in the pre-processing of the models from focus 1 and 2.

The methodology is applicable and useful for the VRPPCdR from focus 1, when the distribution area is large. Different edges of the graph are in different parts of the area. This is considered by different cost coefficients that are corrected for the infrastructural differences. When the area is relatively small, it is more likely that the infrastructural differences are small and neglectable. The application of edge-specific cost coefficients improves the selection of shipments for self-fulfillment: select the ship-to addresses that can be routed via a cheaper infrastructure for self-fulfillment and delegate the addresses in costlier areas to common carriers. In addition, the routing decision is improved because cheaper edges are preferred over shorter but costlier edges.

The methodology is also applicable for the DDP from focus 2. The balance of the estimated tour duration is modelled as a set of constraints. Therefore, infrastructural differences regarding the average expected driving speed could be estimated. This approach would improve the estimator of the expected tour duration and reduce the difference between expected and average tour duration. For example, the driving speed per zip code should be corrected for infrastructural differences. As the infrastructural differences are incorporated, the actual balance of districts improves.

5.2.4 Other approaches to issues of the last mile

Besides the model-based DSS approach to last mile decision problems, there are other approaches for that purpose worthwhile investigating. The perspectives from chapter 1 make many different suggestions and some are outlined in the following sub-sections with no intention of being exhaustive. Table 5.1 proposes basic ideas for the model-based and five other approaches to the issues of the last mile.

5.2.4.1 Structure

The structural approach follows the idea to change the underlying structure (the graph) itself, where the physical flows of goods are managed. The logistics perspective is focused on the process-orientation and thus questions the underlying process map and the supporting systems. For example, the process of loading the CoDi vehicles at the BBP is an object of structural process changes: ex-ante chaotic picking and loading processes could be decoupled such that picking is performed as soon as inbound lines arrive and loading is postponed until every load is re-consolidated. The objective in this example is to speed up the loading in order to depart from the BBP earlier. The SCM perspective considers the physical structure of the whole chain and questions the physical locations not only of the last mile but also of its supply side. For example, the establishment of regional transshipment points can enable faster exchange of loads among network members, such that not all loads need to meet in the central hub. As a result, some lines are shorter and inbound line-haul trucks arrive earlier at the BBP and thus provide more time for the re-consolidation of delivery tour loads. The marketing perspective considers the acquisition of demand. The acquisition of shipments that have suitable origins and destinations changes the sources and the sinks of flows. This purposeful change of the shipment structure enables improved densification of stops and tours (sub-section 1.3.3).

Table 5.1 Approaches to last mile issues

Approach	**Model-based DSS**	**Structure**	**Accounting**
Perspective	Logistics, SCM	Marketing, Logistics, SCM	Logistics, SCM
Object	Decision variables of flows of • Activities • Product quantities • Assignments	• Processual structure • Physical structure • Shipment structure	• Tariffs and discounts • Allocation protocol
Idea	Do everything as-is, but change the values of some variables in order to achieve better objective values.	Change the structure on the last mile in a way that changes the sources, links, and sinks of flows.	Do everything as-is physically, but improve the accuracy, transparency, and fairness of cost accounting and allocations.

Approach	**Market-oriented strategy**	**Technical resources**	**Inter-organizational relations**
Perspective	Marketing, SCM	Logistics	SCM
Object	• Range of service offers • Delivery options • Strategic role of the LSP	• Automatization • Electrification • Human resources	• Horizontal and vertical relations • Cooperative planning • Data sharing
Idea	Identify and conquer services and delivery options that have a great margin.	Do everything as-is, but invest in and employ technical innovations and resources that mitigate the last mile issues sustainably.	Extend the range and intensify the connectedness with long-term partners on the last mile towards more cooperative planning.

5.2.4.2 Accounting

Another approach, which does not make any physical changes, is accounting. The objects are the tariffs that are offered to shipping customers and the transfer prices (allocation protocols) within forwarding alliances. Basic tariffs should be oriented on the fulfillment costs (see sub-section 1.3.3). However, the ring-model builds tariffs based on averages and needs adaptations for some significantly deviating cases. For example, in section 4.5, it is shown that the average costs per km in Cologne are significantly greater than in Regensburg. Therefore, the tariff for shipments to destinations in Cologne should account for higher costs. As a result of that kind of adaptations, the revenues on the last mile would follow the costs. Tariffs may include some discounts for customers who feed large volumes into a logistics network. In section 2, common carriers offer discounts, but there is no discussion and justification of the right value and function of discounts. As said before in sub-section 5.2.2, the tariffs of subcontractors need further research. Rational justification of discounts need to account for the contribution of shipments to the densification of tours and stops. Large volumes of shipments to many dispersed destinations should not be rewarded with discounts because the remaining discounted revenues do not outweigh the marginal cost.

5.2.4.3 Market-oriented strategy

The market-oriented, strategic approach (Porter 2008, 24 et seq.) tries to identify fitting combinations of (a) services, service features, and delivery options and (b) market segments, which generate supra-normal margins. Based on the market's structures (shipments, consigners, destinations, competitors, regulations, innovations), an LSP's position in that market defines where and how to compete in that market. LSPs on the last mile should re-think their own position and their services around the products, which they deliver (Fuller et al. 1993). Classical positions of LSPs differ in their focuses on specific services (what?) or the focus on specific customers (for whom?) (Juga et al. 2008):

(a) the unfocused generalist offering general inventory and transportation services following a strategy of high efficiency
(b) the competence-based specialist (service developer) following a niche strategy and developing specialized services
(c) the 3PL contract LSP who adapts and dedicates its services and assets to few customers' requirements and invests significantly in the relationship with these customers
(d) the customer developer who focuses on dedicated specialty services for particular customers

The market-oriented view assumes that a great fit of the position and the industry structures leads to competitive advantages. For example, a regional rooted and operating LSP, whose stops include many pharmacies, may consider turning into a value-adding LSP specializing on high-frequent, responsive deliveries. Products with very high urgency usually have great margins and thus need responsiveness (speed). Pharmaceutics and spare parts are examples of that kind of products. Specializing on the high-frequent last mile deliveries is thus a worthwhile consideration, if the forces of the market provide this opportunity and the LSP's capabilities are flexible enough.

5.2.4.4 Technical resources

Another approach to the last mile issues is to invest in technical resources and their employment in the company or the network. The idea is that superior resources improve the achievement of goals effectively and lead to supra-normal margins (see Barney 1991 on the resource-based view). The processes of distribution stay mostly as-is, but some tasks within a process are automatized through innovative technology. For example, many sub-processes of consolidation in the BBP, such as picking or sorting, can be automatized.

The electrification of conventional combustion engines contributes to the issue of GHG emissions on the last mile. Last mile distribution takes place where the supply chain touches the customers, where they live and shop. Therefore, electrification of the CoDi fleet is a technical approach that reduces emissions on the last mile. The electric range (distance between recharging) of 100 miles (160km) is not a show stopper for all last mile tours: Pearre et al. (2011) estimate that 32% of the tours in their study can be adapted to meet that electric range. The applicability of electric vehicles is restricted to short delivery tours and it is recommended for dense areas like cities. Due to diesel savings and government subsidies, electric vehicles could well be cost-efficient for the close and small shipments. The new process model for delivery tours with recharging is a topic for further research. This process will fit some type of stops more than others. For example, if stopping, making a delivery including value-added services, and recharging can be done at the same time, then there are some industry-specific use-cases for a fleet of electric CoDis, such as the supply of construction sites.

The deployment of technical innovations also requires investments in qualified personnel. This comprises both the in-house IT department and the pickers, packers, and drivers.

5.2.4.5 Inter-organizational relations

The SCM perspective approaches the inter-organizational relations as a source of competitive advantage. The so-called relational view (Dyer and Singh 1998) assumes that kind of relationships as the source of supra-normal profits. This means, an LSP has to select the right partners (vertical: manufacturers, retailers; horizontal: logistics alliances) and build trusting, mutual depending (Ganesan 1994) relations. The objective of the relational approach is cooperative planning at the interfaces in-between partners. On the one hand, vertical relations smooth the flows of all kinds across stages of the chain. For example, the hand-over of CG at the ramp of the retail store causes costly activities like waiting, screening and counting, paper work, and so on. An agreed inter-organizational process could eliminate these idle times through cooperative planning, early notification of changes, shared data formats, value-adding shelf services, and social collaboration. On the other hand, horizontal relations are expected to produce considerable savings compared to stand-alone LSPs. For example, there is a large number of medium-size forwarding companies in Germany which subcontract an even larger number of small-size carriers. As many LSP operate regionally, those LSPs do not compete. Instead, their regional services complement each other (Cruijssen et al. 2007). Savings are generated through an efficient division of labor in horizontal LSP alliances. The regional complementary services create mutual dependency and transfer pricing is used to allocate those savings in a way that builds trust.

Bibliography

Ahlert, D. (2005). Distributionspolitik. Das Management des Absatzkanals. 3rd ed. Stuttgart: Gustav Fischer (Grundwissen der Ökonomik: Betriebswirtschaftslehre).

Alewell, K. (1968). Ansätze zu einer Konzeption einer ökonomischen Absatztheorie, in: E. Kosiol, E. Sundhoff (Eds.): *Betriebswirtschaft und Marktpolitik.* Beiträge zur Allgemeinen Betriebswirtschaftslehre und zur Betriebswirtschaftlichen Marktlehre. Festschrift für Rudolf Seÿffert zum 75. Geburtstag. Wiesbaden, s.l.: VS Verlag für Sozialwissenschaften, 1–15.

Andler, K. (1929). Rationalisierung der Fabrikation und optimale Losgröße: Oldenbourg.

Anshelevich, E., Dasgupta, A., Kleinberg, J., Tardos, É., Wexler, T., Roughgarden, T. (2008). The Price of Stability for Network Design with Fair Cost Allocation, in: *SIAM Journal on Computing* 38 (4), 1602–1623. DOI: 10.1137/070680096.

Applegate, D. L., Bixby, R. E., Chvatal, V., Cook, W. J. (2011). The Traveling Salesman Problem. A Computational Study. Princeton: Princeton University Press (Princeton series in applied mathematics).

Balas, E. (1989). The prize collecting traveling salesman problem, in: *Networks* 19 (6), 621–636. DOI: 10.1002/net.3230190602.

Balas, E. (2007). The Prize Collecting Traveling Salesman Problem and its Applications, in: G. Gutin, A. P. Punnen (Eds.): *The Traveling Salesman Problem and Its Variations, vol.* 12. New York: Springer Science+Business Media (12), 663–695.

Ball, M. O., Golden, B. L., Assad, A. A., Bodin, L. D. (1983). Planning for truck fleet size in the presence of a common-carrier option, in: *Decision Sciences* 14 (1), 103–120. DOI: 10.1111/j.1540-5915.1983.tb00172.x.

Barney, J. (1991). Firm resources and sustained competitive advantage, in: *Journal of management* 17 (1), 99–120.

Beardwood, J., Halton, J. H., Hammersley, J. M. (1959). The shortest path through many points, in: *Mathematical Proceedings of the Cambridge Philosophical Society, vol.* 55. Cambridge Univ Press, 299–327.

Bell, D. R., Gallino, S., Moreno, A. (2018). The Store is Dead-Long Live the Store, in: *MIT Sloan Management Review* 59 (3), 59–66.

Bienstock, D., Goemans, M., Simchi-Levi, D., Williamson, D. (1993). A note on the prize collecting traveling salesman problem, in: *Mathematical Programming* 59 (1-3), 413–420. DOI: 10.1007/BF01581256.

Bigazzi, A. Y., Figliozzi, M. A. (2012). Impacts of freeway traffic conditions on in-vehicle exposure to ultrafine particulate matter, in: *Atmospheric Environment* 60, 495–503. DOI: 10.1016/j.atmosenv.2012.07.020.

Bigazzi, A. Y., Figliozzi, M. A. (2013). Marginal costs of freeway traffic congestion with on-road pollution exposure externality, in: *Transportation Research Part A: Policy and Practice* 57, 12–24. DOI: 10.1016/j.tra.2013.09.008.

Blumenfeld, D. E., Beckmann, M. J. (1985). Use of continuous space modeling to estimate freight distribution costs, in: *Transportation Research Part A: General* 19 (2), 173–187. DOI: 10.1016/0191-2607(85)90025-1.

Böhnlein, D., Gahm, C., Tuma, A. (2009). A Hybrid Meta-Heuristic for the VRPTW with Cluster-Dependent Tour Starts in the Newspaper Industry, in: IEEE Computer Society (Ed.): *The 42nd Annual Hawaii International Conference on System Sciences (HICSS-42 2009).* 5-8 January 2009, Waikoloa, Big Island, HI, USA, 1–10.

Bolduc, M.-C., Renaud, J., Boctor, F. F. (2007). A Heuristic for the Routing and Carrier Selection Problem, in: *European Journal of Operational Research* 183 (2), 926–932.

Bolduc, M.-C., Renaud, J., Boctor, F. F., Laporte, G. (2008). A Perturbation Metaheuristics for the Vehicle Routing Problem with Private Fleet and Common Carriers, in: *Journal of the Operational Research Society* 59 (6), 776–787.

Box, G. E. P., Hunter, J. S., Hunter, W. G. (2005). Statistics for experimenters. Design, innovation, and discovery. 2nd ed. Hoboken, NJ: Wiley-Interscience (Wiley series in probability and statistics).

Brabänder, C. (2018). Distribution Districting – The Case of In-Night Express Services, in: *Journal of Service Management Research* 2 (4), 33–49.

Bräysy, O. (2003). A Reactive Variable Neighborhood Search for the Vehicle-Routing Problem with Time Windows, in: *INFORMS Journal on Computing* 15 (4), 347–368.

Bretzke, W.-R. (1989). To buy or not to buy. Kriterien für die richtige Entscheidung zwischen Selbsterstellung und Fremdbezug logistischer Leistungen, in: *Internationales Verkehrswesen* 41 (6), 389–395.

Bretzke, W.-R. (2005). Supply chain management: Wege aus einer logistischen Utopie, in: *Logistik Management* 7 (2), 21–30.

Bretzke, W.-R. (2008). Logistische Netzwerke. Berlin, Heidelberg: Springer.

Brown, A. O., Lee, H. L., Petrakian, R. (2000). Xilinx Improves Its Semiconductor Supply Chain Using Product and Process Postponement, in: *Interfaces* 30 (4), 65–80. DOI: 10.1287/inte.30.4.65.11650.

Bundesverbandes Paket und Expresslogistik e. V. (2018). KEP-Studie 2018 – Analyse des Marktes in Deutschland. With assistance of Klaus Esser, Judith Kurte.

Campbell, A., Clarke, L., Kleywegt, A., Savelsbergh, M. (1998). The inventory routing problem, in: *Fleet management and logistics:* Springer, 95–113.

Campbell, J. F. (1990). Freight consolidation and routing with transportation economies of scale, in: *Transportation Research Part B: Methodological* 24 (5), 345–361. DOI: 10.1016/0191-2615(90)90008-M.

Carlsson, J. G. (2012). Dividing a Territory Among Several Vehicles, in: *INFORMS Journal on Computing* 24 (4), 565–577. DOI: 10.1287/ijoc.1110.0479.

Cattaruzza, D., Absi, N., Feillet, D., González-Feliu, J. (2017). Vehicle routing problems for city logistics, in: *EURO Journal on Transportation and Logistics* 6 (1), 51–79. DOI: 10.1007/s13676-014-0074-0.

Chen, C., Pan, S. (2016). Using the Crowd of Taxis to Last Mile Delivery in E-Commerce: a methodological research, in: T. Borangiu, D. Trentesaux, A. Thomas, D. McFarlane (Eds.): *Service Orientation in Holonic and Multi-Agent Manufacturing*. Cham: Springer International Publishing (Studies in Computational Intelligence), 61–70.

Chopra, S. (2003). Designing the distribution network in a supply chain, in: *Transportation Research Part E: Logistics and Transportation Review* 39 (2), 123–140. DOI: 10.1016/S1366-5545(02)00044-3.

Chopra, S., Meindl, P. (2014). Supply Chain Management. Strategie, Planung und Umsetzung. 5., aktualisierte Auflage. Hallbergmoos: Pearson (Always learning).

Christofides, N., Eilon, S. (1969). An Algorithm for the Vehicle-Dispatching Problem, in: *OR* 20 (3), p. 309. DOI: 10.2307/3008733.

Chu, C.-W. (2005). A Heuristic Algorithm for the Truckload and Less-Than-Truckload Problem, in: *European Journal of Operational Research* 165 (3), 657–667.

Clarke, G., Wright, J. W. (1964). Scheduling of Vehicles from a Central Depot to a Number of Delivery Points, in: *Operations research* 12 (4), 568–581.

Converse, P. D. (1954). The other half of marketing, in: *Proceedings of the Twenty-Sixth Boston Conference on Distribution.* Harper and Brothers Publishers Boston, MA, 22–25.

Cooper, L. (1964). Heuristic Methods for Location-Allocation Problems, in: *SIAM Review* 6 (1), 37–53. DOI: 10.1137/1006005.

Cooper, M. C., Lambert, D. M., Pagh, J. D. (1997). Supply Chain Management: More Than a New Name for Logistics, in: *The International Journal of Logistics Management* 8 (1), 1–14. DOI: 10.1108/09574099710805556.

Côté, J.-F., Potvin, J.-Y. (2009). A Tabu Search Heuristic for the Vehicle Routing Problem with Private Fleet and Common Carrier, in: *European Journal of Operational Research* 198 (2), 464–469.

Coughlan, A. T., Anderson, E., Stern, L. W., Ansary, A. I. e. (2006). Marketing channels. 7th ed. Upper Saddle River, NJ: Pearson/Prentice Hall.

Cruijssen, F., Cools, M., Dullaert, W. (2007). Horizontal cooperation in logistics: Opportunities and impediments, in: *Transportation Research Part E: Logistics and Transportation Review* 43 (2), 129–142. DOI: 10.1016/j.tre.2005.09.007.

Daganzo, C. F. (1984). The length of tours in zones of different shapes, in: *Transportation Research Part B: Methodological* 18 (2), 135–145. DOI: 10.1016/0191-2615(84)90027-4.

Daganzo, C. F. (1987). The Break-Bulk Role of Terminals in Many-to-Many Logistic Networks, in: *Operations research* 35 (4), 543–555. DOI: 10.1287/opre.35.4.543.

Dantzig, G. B., Fulkerson, R., Johnson, S. (1954). Solution of a Large-Scale Travelling Salesman Problem, in: *Institute for Operations Research and the Management Sciences* 2 (4), 393–410.

Daskin, M. S. (2010). Service Science. Hoboken, NJ, USA: John Wiley & Sons, Inc.

Defryn, C., Sörensen, K., Cornelissens, T. (2016). The selective vehicle routing problem in a collaborative environment, in: *European Journal of Operational Research* 250 (2), 400–411. DOI: 10.1016/j.ejor.2015.09.059.

Demsetz, H. (1967). Toward a Theory of Property Rights, in: *The American Economic Review* 57 (2), 347–359.

Dettenbach, A., Ubber, S. (2015). Managing Disruptions in Last Mile Distribution, in: *System Sciences (HICSS), 2015 48th Hawaii International Conference on.* IEEE, 1078–1087.

Devari, A., Nikolaev, A. G., He, Q. (2017). Crowdsourcing the last mile delivery of online orders by exploiting the social networks of retail store customers, in: *Transportation Research Part E: Logistics and Transportation Review* 105, 105–122. DOI: 10.1016/j.tre.2017.06.011.

Diaby, M., Ramesh, R. (1995). The Distribution Problem with Carrier Service: A Dual Based Penalty Approach, in: *ORSA Journal on Computing* 7 (1), 24–35. DOI: 10.1287/ijoc.7.1.24.

Drexl, A., Haase, K. (1999). Fast Approximation Methods for Sales Force Deployment, in: *Management Science* 45 (10), 1307–1323. DOI: 10.1287/mnsc.45.10.1307.

Drucker, P. F. (1962). The economy's dark continent, in: *Fortune* 65 (103), 265–270.

Dyer, J. H., Singh, H. (1998). The Relational View: Cooperative Strategy and Sources of Interorganizational Competitive Advantage, in: *Academy of Management Review* 23 (4), 660–679. DOI: 10.5465/amr.1998.1255632.

Edwards, J. B., McKinnon, A. C., Cullinane, S. L. (2010). Comparative analysis of the carbon footprints of conventional and online retailing, in: *International Journal of Physical Distribution & Logistics Management* 40 (1/2), 103–123. DOI: 10.1108/09600031011018055.

Eisele, W. L., Schrank, D. L. (2010). Conceptual Framework and Trucking Application for Estimating Impact of Congestion on Freight, in: *Transportation Research Record: Journal of the Transportation Research Board* 2168 (1), 94–103. DOI: 10.3141/2168-12.

Feige, D., Klaus, P. (2008). Modellbasierte Entscheidungsunterstützung in der Logistik. mit EUS-Lehrsoftware auf CD-ROM. Hamburg: DVV Media Group, Dt. Verkehrs-Verl. (Edition Logistik, 10).

Figliozzi, M. A. (2010). The impacts of congestion on commercial vehicle tour characteristics and costs, in: *Transportation Research Part E: Logistics and Transportation Review* 46 (4), 496–506. DOI: 10.1016/j.tre.2009.04.005.

Fink, A., Kliewer, N., Mattfeld, D., Mönch, L., Rothlauf, F., Schryen, G., Suhl, L., Voß, S. (2014). Modellbasierte Entscheidungsunterstützung in Produktions- und Dienstleistungsnetzwerken, in: *WIRTSCHAFTSINFORMATIK* 56 (1), 21–29. DOI: 10.1007/s11576-013-0402-2.

Fischer, M. E. (2001). "Available-to-promise": Aufgaben u. Verfahren im Rahmen d. supply chain management.

Fleischmann, B. (1998). Design of freight traffic networks, in: B. Fleischmann, J.A.E.E. van Nunen, M. Grazia Speranza, P. Stähly (Eds.): *Advances in distribution logistics*. Berlin, Heidelberg, New York: Springer, 55–81.

Fleischmann, B., Meyr, H., Wagner, M. (2008). Advanced Planning, in: H. Stadtler, C. Kilger (Eds.): *Supply Chain Management and Advanced Planning.* Concepts, Models, Software, and Case Studies. 4th ed. Berlin: Springer, 81–106.

Fleischmann, B., Paraschis, J. N. (1988). Solving a large scale districting problem. A case report, in: *Computers & Operations Research* 15 (6), 521–533. DOI: 10.1016/0305-0548(88)90048-2.

Franceschetti, A., Jabali, O., Laporte, G. (2017). Continuous approximation models in freight distribution management, in: *TOP* 25 (3), 413–433. DOI: 10.1007/s11750-017-0456-1.

Frankfort-Nachmias, C., Nachmias, D. (2008). Research methods in the social sciences. 7th ed. New York: Worth Publishers.

Frerich, J., Müller, G. (2004). Europäische Verkehrspolitik. Von den Anfängen bis zur Osterweiterung der Europäischen Union. München: Oldenbourg Verlag.

Fuller, J. B., O'Conor, J., Rawlinson, R. (1993). Tailored logistics: The next advantage, in: *Harvard Business Review* 71 (3), 87–98.

Gahm, C., Brabänder, C., Tuma, A. (2017). Vehicle routing with private fleet, multiple common carriers offering volume discounts, and rental options, in: *Transportation Research Part E: Logistics and Transportation Review* 97, 192–216. DOI: 10.1016/j.tre.2016.10.010.

Gahm, C., Dünnwald, B., Sahamie, R. (2014). A multi-criteria master production scheduling approach for special purpose machinery, in: *International Journal of Production Economics* 149, 89–101. DOI: 10.1016/j.ijpe.2013.10.011.

Ganesan, S. (1994). Determinants of Long-Term Orientation in Buyer-Seller Relationships, in: *Journal of Marketing* 58 (2), 1–19. DOI: 10.2307/1252265.

Gelders, L., Cattrysse, D. (1991). Public Waste Collection. A Case Study, in: *JORBEL* 31 (1).

Golden, B. L., Assad, A. A. (Eds.) (1988). Vehicle Routing. Methods and Studies. Amsterdam (Studies in Management Science and Systems, 16).

Golden, B. L., Wasil, E. A., Kelly, J. P., Chao, I.-M. (1998). The Impact of Metaheuristics on Solving the Vehicle Routing Problem: Algorithms, Problem Sets, and Computational Results, in: T. Gabriel Crainic, G. Laporte (Eds.): *Fleet Management and Logistics*. New York: Springer Science+Business Media, 33–56.

Grefenstette, J. J. (2000). Proportional selection and sampling algorithms, in: T. Bäck, D. B. Fogel, Z. Michalewicz (Eds.): *Evolutionary computation*. New York, London: Taylor & Francis, 172–180.

Hall, R. W. (1991). Route selection on freight networks with weight and volume constraints, in: *Transportation Research Part B: Methodological* 25 (4), 175–189. DOI: 10.1016/0191-2615(91)90002-Z.

Hall, R. W., Racer, M. (1995). Transportation with common carrier and private fleets: system assignment and shipment frequency optimization, in: *IIE Transactions* 27 (2), 217–225. DOI: 10.1080/07408179508936734.

Hansen, H. R., Neumann, G., Mendling, J. (2015). Wirtschaftsinformatik. Grundlagen und Anwendungen. 11th ed. Berlin, München, Boston: De Gruyter.

Hansen, P., Mladenović, N. (2001). Variable neighborhood search: Principles and applications, in: *European Journal of Operational Research* 130 (3), 449–467. DOI: 10.1016/S0377-2217(00)00100-4.

Hansen, P., Mladenović, N., Moreno Perez, J. A. (2010). Variable neighbourhood search: methods and applications, in: *Annals of Operations Research* 175 (1), 367–407. DOI: 10.1007/s10479-009-0657-6.

Haugland, D., Ho, S. C., Laporte, G. (2007). Designing delivery districts for the vehicle routing problem with stochastic demands, in: *European Journal of Operational Research* 180 (3), 997–1010. DOI: 10.1016/j.ejor.2005.11.070.

Helsgaun, K. (2000). An Effective Implementation of the Lin-Kernighan Traveling Salesman Heuristic, in: *European Journal of Operational Research* 126 (1), 106–130.

Helsgaun, K. (2009). General K-Opt Submoves for the Lin–Kernighan TSP Heuristic, in: *Mathematical Programming Computation* 1 (2-3), 119–163.

Hemmelmayr, V. C., Doerner, K. F., Hartl, R. F. (2009). A Variable Neighborhood Search Heuristic for Periodic Routing Problems, in: *European Journal of Operational Research* 195 (3), 791–802.

Hertz, A., Mittaz, M. (2001). A Variable Neighborhood Descent Algorithm for the Undirected Capacitated Arc Routing Problem, in: *Transportation Science* 35 (4), 425–434.

Hitch, C. (1957). Letter to the Editor—Operations Research and National Planning—A Dissent, in: *Operations research* 5 (5), 718–723. DOI: 10.1287/opre.5.5.718.

Hotz, A., Fost, M. (2016). Die „Amazonisierung" des Konsums – Game-Changer Amazon, in: D. Schallmo, A. Rusnjak, J. Anzengruber, T. Werani, M. Jünger (Eds.): *Digitale Transformation von Geschäftsmodellen: Grundlagen, Instrumente und Best Practices*. Wiesbaden: Springer Science and Business Media; Springer Gabler, 669–696.

Houlihan, J. B. (1985). International Supply Chain Management, in: *International Journal of Physical Distribution & Materials Management* 15 (1), 22–38. DOI: 10.1108/eb014601.

Ihde, G. B. (2001). Transport, Verkehr, Logistik. Gesamtwirtschafliche Aspekte und einzelwirtschaftliche Handhabung. 3rd ed. München: Vahlen (Vahlens Handbücher der Wirtschafts- und Sozialwissenschaften).

Inderfurth, K., Minner, S. (1998). Safety stocks in multi-stage inventory systems under different service measures, in: *European Journal of Operational Research* 106 (1), 57–73. DOI: 10.1016/S0377-2217(98)00210-0.

Jafari, H., Nyberg, A., Hilletofth, P. (2016). Postponement and logistics flexibility in retailing. A multiple case study from Sweden, in: *Industrial Management & Data Systems* 116 (3), 445–465. DOI: 10.1108/IMDS-06-2015-0257.

Jharkharia, S., Shankar, R. (2007). Selection of logistics service provider: An analytic network process (ANP) approach, in: *Omega* 35 (3), 274–289. DOI: 10.1016/j.omega.2005.06.005.

Johnson, D. S., McGeoch, L. A. (2007). Experimental Analysis of Heuristics for the STSP, in: G. Gutin, A. P. Punnen (Eds.): *The Traveling Salesman Problem and Its Variations*. New York: Springer Science+Business Media (12), 369–443.

Juga, J., Pekkarinen, S., Kilpala, H. (2008). Strategic positioning of logistics service providers, in: *International Journal of Logistics Research and Applications* 11 (6), 443–455. DOI: 10.1080/13675560802373472.

Kalcsics, J. (2015). Districting Problems, in: G. Laporte (Ed.): *Location science*. Cham: Springer, 595–622.

Karapetyan, D., Gutin, G. (2011). Lin–Kernighan heuristic adaptations for the generalized traveling salesman problem, in: *European Journal of Operational Research* 208 (3), 221–232. DOI: 10.1016/j.ejor.2010.08.011.

Kellner, F. (2016a). Allocating greenhouse gas emissions to shipments in road freight transportation: Suggestions for a global carbon accounting standard, in: *Energy Policy* 98, 565–575. DOI: 10.1016/j.enpol.2016.09.030.

Kellner, F. (2016b). Exploring the impact of traffic congestion on CO2 emissions in freight distribution networks, in: *Logistics Research* 9 (1), p. 347. DOI: 10.1007/s12159-016-0148-5.

Kellner, F., Otto, A., Brabänder, C. (2017). Bringing infrastructure into pricing in road freight transportation – A measuring concept based on navigation service data, in: *Transportation Research Procedia* 25, 794–805. DOI: 10.1016/j.trpro.2017.05.458.

Kellner, F., Otto, A., Busch, A. (2013). Understanding the robustness of optimal FMCG distribution networks, in: *Logistics Research* 6 (4), 173–185. DOI: 10.1007/s12159-012-0097-6.

Kirsch, W. (1998). Betriebswirtschaftslehre. Eine Annäherung aus der Perspektive der Unternehmensführung. 5th ed. München: Kirsch (Münchener Schriften zur angewandten Führungslehre, 70).

Klaus, P. (2002). Die dritte Bedeutung der Logistik. Beiträge zur Evolution logistischen Denkens. Hamburg: Dt. Verkehrs-Verl. (Edition Logistik, 1).

Klaus, P., Sheffi, Y. (2012). Intermediate Commentary: On the Evolution of Three Meanings of Logistics, in: P. Klaus, S. Müller (Eds.): *The Roots of Logistics*. Berlin, Heidelberg: Springer Berlin Heidelberg, 117–126.

Klincewicz, J. G., Luss, H., Pilcher, M. G. (1990). Fleet Size Planning when Outside Carrier Services Are Available, in: *Transportation Science* 24 (3), 169–182.

Koether, R. (2014). Distributionslogistik. Effiziente Absicherung der Lieferfähigkeit. 2nd ed. Wiesbaden: Springer Gabler.

Kopfer, H., Wang, X. (2009). Combining Vehicle Routing with Forwarding - Extension of the Vehicle Routing Problem by Different Types of Sub-contraction, in: *Journal of the Korean Institute of Industrial Engineers* 35 (1), 1–14.

Krajewska, M. A., Kopfer, H. (2009). Transportation Planning in Freight Forwarding Companies. Tabu Search Algorithm for the Integrated Operational Transportation Planning Problem, in: *European Journal of Operational Research* 197 (2), 741–751.

Kuhn, H., Holzapfel, A., Ostermeier, M. (2018). Handelslogistik, in: H. Corsten, R. Gössinger, T. S. Spengler, H. Corsten, R. Gössinger, T. S. Spengler (Eds.): *Handbuch Produktions- und Logistikmanagement in Wertschöpfungsnetzwerken*. Berlin, Boston: De Gruyter, 716–736.

Kumar, K. (2001). Technology for supporting supply chain management: introduction, in: *Communications of the ACM* 44 (6), 58–61. DOI: 10.1145/376134.376165.

Kunkel, M., Schwind, M. (2012). Vehicle Routing with Driver Learning for Real World CEP Problems, in: *45th Hawaii International Conference on System Sciences (HICSS).* Maui, HI, USA, 1315–1322.

Kytöjoki, J., Nuortio, T., Bräysy, O., Gendreau, M. (2007). An Efficient Variable Neighborhood Search Heuristic for Very Large Scale Vehicle Routing Problems, in: *Computers & Operations Research* 34 (9), 2743–2757.

Lambert, D. M. (Ed.) (2014). Supply chain management. Processes, partnerships, performance. Supply Chain Management Institute. 4th ed. Ponte Vedra Beach.

Lambert, D. M., Burduroglu, R. (2000). Measuring and Selling the Value of Logistics, in: *The International Journal of Logistics Management* 11 (1), 1–18. DOI: 10.1108/09574090010806038.

Larke, R., Kilgour, M., O'Connor, H. (2018). Build touchpoints and they will come: transitioning to omnichannel retailing, in: *International Journal of Physical Distribution & Logistics Management* 48 (4), 465–483. DOI: 10.1108/IJPDLM-09-2016-0276.

Lee, H. L., Padmanabhan, V., Whang, S. (1997). The bullwhip effect in supply chains, in: *Sloan management review* 38, 93–102.

Lee, H. L., Whang, S. (2001). Winning the last mile of e-commerce, in: *MIT Sloan Management Review* 42 (4), 54–62.

Lei, H., Laporte, G., Guo, B. (2012). Districting for routing with stochastic customers, in: *EURO Journal on Transportation and Logistics* 1 (1), 67–85. DOI: 10.1007/s13676-012-0005-x.

Lenstra, J. K., Rinnooy Kan, A. H. G. (1981). Complexity of vehicle routing and scheduling problems, in: *Networks* 11 (2), 221–227. DOI: 10.1002/net.3230110211.

Levy, J. I., Buonocore, J. J., Stackelberg, K. von (2010). Evaluation of the public health impacts of traffic congestion: a health risk assessment, in: *Environmental Health* 9 (65). DOI: 10.1186/1476-069X-9-65.

Lin, S. (1965). Computer Solutions of the Traveling Salesman Problem, in: *Bell System Technical Journal* 44 (10), 2245–2269.

Lin, S., Kernighan, B. W. (1973). An Effective Heuristic Algorithm for the Traveling-Salesman Problem, in: *Operations research* 21 (2), 498–516.

Lindblom, C. E. (1959). The Science of "Muddling Through", in: *Public Administration Review* 19 (2), 79–88.

Liu, R., Jiang, Z., Liu, X., Chen, F. (2010). Task selection and routing problems in collaborative truckload transportation, in: *Transportation Research Part E: Logistics and Transportation Review* 46 (6), 1071–1085. DOI: 10.1016/j.tre.2010.05.003.

Lomax, T. J., Schrank, D. L., Turner, S., Geng, L., Li, Y., Koncz, N. (2011). Real-Timing the 2010 Urban Mobility Report. Washington, DC. Available online at http://utcm.tamu.edu/publications/final_reports/Lomax_10-65-55.pdf, checked on 12/10/2014.

Macioszek, E. (2018). First and Last Mile Delivery – Problems and Issues, in: G. Sierpiński (Ed.): *Advanced Solutions of Transport Systems for Growing Mobility.* 14th Scientific and Technical Conference "Transport Systems. Theory & Practice 2017" Selected Papers, vol. 631. Cham: Springer International Publishing; Imprint: Springer (Advances in Intelligent Systems and Computing, 631), 147–154.

Magee, J. F. (1960). The Logistics of Distribution, in: *Harvard Business Review* 38 (4), 89–101.

Menon, M. K., McGinnis, M. A., Ackerman, K. B. (1998). Selection criteria for providers of third-party logistics services: An exploratory study, in: *Journal of Business Logistics* 19 (1), 121–137.

Meyer, C. M., Kopfer, H. (2008). Restrictions for the Operational Transportation Planning by Regulations on Drivers' Working Hours, in: A. Bortfeldt, J. Homberger, H. Kopfer, G. Pankratz, R. Strangmeier (Eds.): *Intelligent Decision Support.* Current Challenges and Approaches. Wiesbaden: Gabler, 177–186.

Miller, C. E., Tucker, A. W., Zemlin, R. A. (1960). Integer programming formulations and traveling salesman problems, in: *Journal of Association for Computing Machinery*, 326–329.

Mladenović, N., Hansen, P. (1997). Variable Neighborhood Search, in: *Computers & Operations Research* 24 (11), 1097–1100.

Mole, R. H. (1979). A Survey of Local Delivery Vehicle Routing Methodology, in: *The Journal of the Operational Research Society* 30 (3), p. 245. DOI: 10.2307/3009605.

Morgan, M. S. (2012). The world in the model. How economists work and think. Cambridge, New York: Cambridge University Press.

Müller, S., Klaus, P. (2009). More Expensive or Too Expensive? Calculating Delivery Costs in Europe, in: R. Shariatmadari, F. Josef Schoppengerd, A. Otto (Eds.): *Direct Store Delivery*. Berlin, Heidelberg: Springer, 145–160.

Nagata, Y., Bräysy, O. (2009). Edge assembly-based memetic algorithm for the capacitated vehicle routing problem, in: *Networks* 54 (4), 205–215. DOI: 10.1002/net.20333.

Nguyen, H. N., Rainwater, C. E., Mason, S. J., Pohl, E. A. (2014). Quantity discount with freight consolidation, in: *Transportation Research Part E: Logistics and Transportation Review* 66, 66–82. DOI: 10.1016/j.tre.2014.03.004.

Niemi, R. G., Grofman, B., Carlucci, C., Hofeller, T. (1990). Measuring Compactness and the Role of a Compactness Standard in a Test for Partisan and Racial Gerrymandering, in: *The Journal of Politics* 52 (4), 1155–1181. DOI: 10.2307/2131686.

Novaes, A. G.N., Souza de Cursi, J. E., da Silva, A. C.L., Souza, J. C. (2009). Solving continuous location–districting problems with Voronoi diagrams, in: *Computers & Operations Research* 36 (1), 40–59. DOI: 10.1016/j.cor.2007.07.004.

O'Brien, H. (1975). Distribution models and transport planning, in: *Retail and Distribution Management* 3 (4), 49–52. DOI: 10.1108/eb060374.

Osman, I. H. (1993). Metastrategy Simulated Annealing and Tabu Search Algorithms for the Vehicle Routing Problem, in: *Annals of Operations Research* 41 (4), 421–451.

Otto, A. (2002). Management und Controlling von Supply Chains. Ein Modell auf der Basis der Netzwerktheorie. Wiesbaden: Deutscher Universitätsverlag (Neue Betriebswirtschaftliche Forschung, 290).

Otto, A. (2008). Auftragsabwicklung, in: P. Klaus, W. Krieger (Eds.): *Gabler-Lexikon Logistik.* Management logistischer Netzwerke und Flüsse. 4th ed. Wiesbaden: Gabler, 19–25.

Otto, A., Kotzab, H. (2003). Does supply chain management really pay? Six perspectives to measure the performance of managing a supply chain, in: *European Journal of Operational Research* 144 (2), 306–320. DOI: 10.1016/S0377-2217(02)00396-X.

Otto, A., Lukesch, M., Brabänder, C., Kellner, F. (2018). Konsumgüterdistribution, in: H. Corsten, R. Gössinger, T. S. Spengler (Eds.): *Handbuch Produktions- und Logistikmanagement in Wertschöpfungsnetzwerken*. Berlin, Boston: De Gruyter, 737–758.

Özkaya, E., Keskinocak, P., Roshan Joseph, V., Weight, R. (2010). Estimating and benchmarking Less-than-Truckload market rates, in: *Transportation Research Part E: Logistics and Transportation Review* 46 (5), 667–682. DOI: 10.1016/j.tre.2009.09.004.

Pagh, J. D., Cooper, M. C. (1998). Supply chain postponement and speculation strategies. How to choose the right strategy, in: *Journal of Business Logistics* 19 (2), 13–33.

Parry, I. W.H. (2008). Pricing Urban Congestion. Washington, DC. Available online at http://www.rff.org/rff/Documents/RFF-DP-08-35.pdf, checked on 12/7/2014.

Pearre, N. S., Kempton, W., Guensler, R. L., Elango, V. V. (2011). Electric vehicles: How much range is required for a day’s driving?, in: *Transportation Research Part C: Emerging Technologies* 19 (6), 1171–1184. DOI: 10.1016/j.trc.2010.12.010.

Picot, A., Reichwald, R., Wigand, R. T. (2010). Die grenzenlose Unternehmung. Information, Organisation und Management; Lehrbuch zur Unternehmensführung im Informationszeitalter. 5th ed. Wiesbaden: Gabler.

Plastria, F. (2011). The Weiszfeld Algorithm: Proof, Amendments, and Extensions, in: H. A. Eiselt, V. Marianov (Eds.): *Foundations of Location Analysis, vol.* 155. Boston, MA: Springer US (International series in operations research & management science), 357–389.

Plowman, E. G. (1964). Lectures on elements of business logistics. Stanford (Stanford transportation series.).

Polacek, M., Hartl, R. F., Doerner, K. F. (2004). A Variable Neighborhood Search for the Multi Depot Vehicle Routing Problem with Time Windows, in: *Journal of Heuristics* 10 (6), 613–627.

Porter, M. E. (2008). Wettbewerbsstrategie. Methoden zur Analyse von Branchen und Konkurrenten. 11th ed. Frankfurt am Main: Campus Verlag.

Potvin, J.-Y., Naud, M.-A. (2011). Tabu Search with Ejection Chains for the Vehicle Routing Problem with Private Fleet and Common Carrier, in: *Journal of Operational Research Society* 62 (2), 326–336.

Pratchett, T., Stewart, I., Cohen, J. S. (2013). The Science of Discworld. Rev. ed. London: Ebury Press.

Public Policy Institute of California (PPIC) (2002). PPIC Statewide Survey: Special Survey on Land Use. San Francisco, CA. Available online at http://www.ppic.org/content/pubs/survey/S_1102MBS.pdf, checked on 12/4/2014.

Punakivi, M., Yrjölä, H., Holmström, J. (2001). Solving the last mile issue: reception box or delivery box?, in: *International Journal of Physical Distribution & Logistics Management* 31 (6), 427–439. DOI: 10.1108/09600030110399423.

Raffée, H. (1974). Betriebswirtschaftslehre im Grundstudium der Wirtschaftswissenschaft. Göttingen: Vandenhoeck & Ruprecht (UTB Uni-Taschenbücher, 97).

Rao, A. M., Rao, K. R. (2012). Measuring urban traffic congestion - a review, in: *International Journal for Traffic and Transport Engineering* 2 (4), 286–305. DOI: 10.7708/ijtte.2012.2(4).01.

Reichwald, R., Dietel, B. (1991). Produktionswirtschaft, in: E. Heinen, A. Picot (Eds.): *Industriebetriebslehre.* Entscheidungen im Industriebetrieb. 9th ed. Wiesbaden: Gabler Verlag, 395–622.

ReVelle, C. S., Eiselt, H. A. (2005). Location analysis: A synthesis and survey, in: *European Journal of Operational Research* 165 (1), 1–19. DOI: 10.1016/j.ejor.2003.11.032.

Rieksts, B. Q., Ventura, J. A. (2008). Optimal inventory policies with two modes of freight transportation, in: *European Journal of Operational Research* 186 (2), 576–585. DOI: 10.1016/j.ejor.2007.01.042.

Rodrigues, A. M., Ferreira, J. S. (2015). Sectors and routes in solid waste collection, in: *Operational Research:* Springer, 353–375.

Rouquet, A., Vauché, L. (2015). A Typology of Logistics Pooling in Supply Chains, in: *Supply Chain Forum: An International Journal* 16 (2), 2–12. DOI: 10.1080/16258312.2015.11673825.

Sarzynski, A., Wolman, H., Galster, G., Hanson, R. (2006). Testing the conventional wisdom about land use and traffic congestion: The more we sprawl, the less we move?, in: *Urban Studies* 43 (3), 601–626. DOI: 10.1080/00420980500452441.

Savelsbergh, M. W. P., Goetschalckx, M. (1995). A comparison of the efficiency of fixed versus variable vehicle routes, in: *Journal of Business Logistics* 16 (1), 163–188.

Schmidthöfer, A. (2004). Taktische Bestimmung von Tourgebieten für die speditionelle Tourenplanung. Univ., Disseration Augsburg, 2003. Aachen: Shaker Verlag (Berichte aus der Betriebswirtschaft).

Schneeweiß, C., Kühn, M. (1990). Zur Definition und gegenseitigen Abgrenzung der Begriffe Flexibilität, Elastizität und Robustheit, in: *Schmalenbachs Zeitschrift für betriebswirtschaftliche Forschung: Zfbf* 42 (5).

Schneider, M. (2016). The vehicle-routing problem with time windows and driver-specific times, in: *European Journal of Operational Research* 250 (1), 101–119. DOI: 10.1016/j.ejor.2015.09.015.

Schögel, M. (2012). Distributionsmanagement. Das Management der Absatzkanäle. München: Vahlen (Vahlens Handbücher der Wirtschafts- und Sozialwissenschaften).

Schrank, D. L., Lomax, T. J. (2002). The 2002 urban mobility report. Texas Transportation Institute. Available online at http://www.dart.org/news/TTIUrbanMobility2002.pdf, checked on 12/4/2014.

Schwemmer, M. (2016). Top 100 der Logistik. Marktgrößen, Marktsegmente und Marktführer; eine Studie der Fraunhofer Arbeitsgruppe für Supply Chain Services SCS. 2016/2017. Hamburg: DVV Media Group GmbH.

Sheen, G.-J., Tsao, Y.-C. (2007). Channel coordination, trade credit and quantity discounts for freight cost, in: *Transportation Research Part E: Logistics and Transportation Review* 43 (2), 112–128. DOI: 10.1016/j.tre.2005.08.001.

Shirabe, T. (2009). Districting Modeling with Exact Contiguity Constraints, in: *Environment and Planning B: Planning and Design* 36 (6), 1053–1066. DOI: 10.1068/b34104.

Simpson, K. F. (1958). In-Process Inventories, in: *Operations research* 6 (6), 863–873. DOI: 10.1287/opre.6.6.863.

Specht, G., Fritz, W. (2005). Distributionsmanagement. 4th ed. Stuttgart: Kohlhammer (Kohlhammer Edition Marketing).

Stachowiak, H. (1973). Allgemeine Modelltheorie. Wien: Springer.

Stenger, A., Schneider, M., Goeke, D. (2013a). The prize-collecting vehicle routing problem with single and multiple depots and non-linear cost, in: *EURO Journal on Transportation and Logistics* 2 (1-2), 57–87. DOI: 10.1007/s13676-013-0022-4.

Stenger, A., Vigo, D., Enz, S., Schwind, M. (2013b). An Adaptive Variable Neighborhood Search Algorithm for a Vehicle Routing Problem Arising in Small Package Shipping, in: *Transportation Science* 47 (1), 64–84. DOI: 10.1287/trsc.1110.0396.

Stüber, E., Heinick, H., Leyendecker, C. (2018). Amazonisierung des Konsums. IFH Köln. Köln.

Sun, L., Karwan, M. H., Gemici-Ozkan, B., Pinto, J. M. (2015). Estimating the long-term cost to serve new customers in joint distribution, in: *Computers & Industrial Engineering* 80, 1–11. DOI: 10.1016/j.cie.2014.11.012.

Taylor, J., Casavant, K., Moore, D., Sage, J., Ivanov, B. (2012). The Economic Impact of Increased Congestion for Freight Dependent Businesses in Washington State. Technical Report. Washington, DC. Available online at http://www.wsdot.wa.gov/NR/rdonlyres/0DA2A843-8BC3-41B7-A0F3-C72A610BEA90/0/EconomicImpactCongestion.pdf, checked on 3/19/2015.

Tempelmeier, H. (2015). Bestandsmanagement in Supply Chains. 5th ed. Norderstedt: BoD - Books on Demand.

Tersine, R. J., Barman, S. (1991). Economic Inventory/Transport Lot, Sizing with Quantity and Freight Rate Discounts, in: *Decision Sciences* 22 (5), 1171–1179. DOI: 10.1111/j.1540-5915.1991.tb01914.x.

Thommen, J.-P., Achleitner, A.-K. (2006). Allgemeine Betriebswirtschaftslehre. Umfassende Einführung aus managementorientierter Sicht. 5th ed. Wiesbaden: Gabler.

TomTom (2014). TomTom European Traffic Index. Available online at http://www.tomtom.com/lib/doc/pdf/2014-05-14%20TomTomTrafficIndex2013annualEur-km.pdf, checked on 12/4/2014.

Toptal, A., Bingöl, S. O. (2011). Transportation pricing of a truckload carrier, in: *European Journal of Operational Research* 214 (3), 559–567. DOI: 10.1016/j.ejor.2011.05.005.

Tsao, Y.-C., Lu, J.-C. (2012). A supply chain network design considering transportation cost discounts, in: *Transportation Research Part E: Logistics and Transportation Review* 48 (2), 401–414. DOI: 10.1016/j.tre.2011.10.004.

Turkensteen, M., Klose, A. (2012). Demand dispersion and logistics costs in one-to-many distribution systems, in: *European Journal of Operational Research* 223 (2), 499–507. DOI: 10.1016/j.ejor.2012.06.008.

Tyan, J., Wee, H.-M. (2003). Vendor managed inventory: a survey of the Taiwanese grocery industry, in: *Journal of Purchasing and Supply Management* 9 (1), 11–18. DOI: 10.1016/S0969-7012(02)00032-1.

Ulrich, H. (1968). Die Unternehmung als produktives soziales System. Bern, Stuttgart: Paul Haupt (Unternehmung und Unternehmungsführung, 1).

Umundum, P. (2015). Paradigmenwechsel auf der letzten Meile, in: P. H. Voß (Ed.): *Logistik - eine Industrie, die (sich) bewegt.* Strategien und Lösungen entlang der Supply Chain 4.0: Springer Gabler, 119–134.

Upton, D. M. (1994). The management of manufacturing flexibility, in: *California management review* 36 (2), p. 72.

Vahrenkamp, R. (2011). The Calculation of Costs for Delivery Routes. Edited by Universität Kassel. Kassel (No. 71/2011). Available online at http://www.vahrenkamp.org/Nr71_Cost_Delivery.pdf, checked on 11/27/2018.

Vakulenko, Y., Hellström, D., Hjort, K. (2018). What's in the parcel locker? Exploring customer value in e-commerce last mile delivery, in: *Journal of Business Research* 88, 421–427. DOI: 10.1016/j.jbusres.2017.11.033.

van Oudheusden, D., Cattrysse, D., Lotan, T. (1999). On the importance of districting and its potential impact on routing, in: *World Transport Research: Selected Proceedings of the 8th World Conference on Transport Research*, 521–531.

van Ryzin, G. J. (2001). Analyzing Inventory Cost and Service in Supply Chains, in: *Columbia Business School Press*.

Vidal, T., Maculan, N., Ochi, L. S., Vaz Penna, P. H. (2015). Large Neighborhoods with Implicit Customer Selection for Vehicle Routing Problems with Profits, in: *Transportation Science* Articles in Advance, 1–15. DOI: 10.1287/trsc.2015.0584.

Volgenant, T., Jonker, R. (1987). On Some Generalizations of the Travelling-Salesman Problem, in: *The Journal of the Operational Research Society* 38 (11), 1073–1079. DOI: 10.2307/2582232.

Vollmann, T. E., Berry, W. L., Whybark, D. C., Jacobs, F. R. (2005). Manufacturing Planning and Control Systems for Supply Chain Management. 5th ed. New York, Chicago, San Francisco: McGraw-Hill.

Wang, X., Kopfer, H. (2014). Collaborative transportation planning of less-than-truckload freight, in: *OR Spectrum* 36 (2), 357–380. DOI: 10.1007/s00291-013-0331-x.

Wang, X., Kopfer, H., Gendreau, M. (2014). Operational transportation planning of freight forwarding companies in horizontal coalitions, in: *European Journal of Operational Research* 237, 1133–1141. DOI: 10.1016/j.ejor.2014.02.056.

Wener, R., Evans, G., Boately, P. (2005). Commuting Stress: Psychophysiological Effects of a Trip and Spillover into the Workplace, in: *Transportation Research Record* 1924 (1), 112–117. DOI: 10.3141/1924-14.

Weng, K. Z. (1995). Modeling quantity discounts under general price-sensitive demand functions. Optimal policies and relationships, in: *European Journal of Operational Research* 86 (2), 300–314. DOI: 10.1016/0377-2217(94)00104-K.

Wilcox, J. B., Howell, R. D., Kuzdrall, P., Britney, R. (1987). Price Quantity Discounts. Some Implications for Buyers and Sellers, in: *Journal of Marketing* 51 (3), p. 60. DOI: 10.2307/1251647.

Wittenbrink, P. (2011). Transportkostenmanagement im Straßengüterverkehr. Grundlagen - Optimierungspotenziale - Green Logistic. 1st ed. Wiesbaden: Gabler.

Wong, K. F., Beasley, J. E. (1984). Vehicle routing using fixed delivery areas, in: *Omega* 12 (6), 591–600. DOI: 10.1016/0305-0483(84)90062-8.

Wong, R. (2008). Vehicle Routing for Small Package Delivery and Pickup Services, in: R. Sharda, S. Voß, B. Golden, S. Raghavan, E. Wasil (Eds.): *The Vehicle Routing Problem: Latest Advances and New Challenges, vol.* 43. 1st ed.: Springer (Operations research / computer science interfaces series, 43), 475–485.

Yaman, H. (2006). Formulations and Valid Inequalities for the Heterogeneous Vehicle Routing Problem, in: *Mathematical Programming* 106 (2), 365–390.

Zhu, W., Erikstad, S. O., Nowark, M. P. (2014). Emission allocation problems in the maritime logistics chain, in: *EURO Journal on Transportation and Logistics* 3 (1), 35–54. DOI: 10.1007/s13676-013-0029-x.

Ziebuhr, M., Kopfer, H. (2014). The Integrated Operational Transportation Planning Problem with Compulsory Requests, in: R. G. González-Ramírez, F. Schulte, S. Voß, J. A. Ceroni Díaz (Eds.): *Computational Logistics, vol.* 8760. 5th International Conference, ICCL 2014. Valparaiso, Chile, September 24-26. Cham, Heidelberg, NewYork, Dordrecht, London: Springer International Publishing (Lecture Notes in Computer Science), 1–15.