AGENT-BASED MODEL FOR THE IN- AND OUT-GOING PARCEL QUANTITIES OF COMPANIES IN AN URBAN AREA

Lukas Barthelmes
Tel: +49 721 608-44119, Fax: +49 721 608-46777, Email: lukas.barthelmes@kit.edu
https://orcid.org/0000-0003-1122-853X

Mehmet Emre Görgülü
Tel: +49 721 608-43048, Fax: +49 721 608-46777, Email: mehmet.goerguelue@kit.edu
https://orcid.org/0000-0001-5845-5073

Jelle Kübler
Tel: +49 721 608-42256, Fax: +49 721 608-46777, Email: jelle.kuebler@kit.edu
https://orcid.org/0000-0001-8095-1815

Martin Kagerbauer
Tel: +49 721 608-47734, Fax: +49 721 608-46777, Email: martin.kagerbauer@kit.edu
https://orcid.org/0000-0003-4252-7874

Peter Vortisch
Tel: +49 721 608-42255, Fax: +49 721 608-46777, Email: peter.vortisch@kit.edu
https://orcid.org/0000-0003-1647-2435

Address
Karlsruhe Institute of Technology
Institute for Transport Studies
Kaiserstrasse 12, D-76131 Karlsruhe, Germany

Word Count: 7203 words + 1 table (250 words per table) = 7453 words

Submission Date: March 13, 2023

ABSTRACT

Increasing e-commerce activities, accelerated by the COVID-19 pandemic, have led to a substantial growth of parcel volumes that must be delivered and picked up by courier, express, and parcel service providers (CEPSPs). In transportation planning, freight demand models are used to evaluate the effects of such developments. Even though the B2B sector contributes considerably to parcel volumes, existing models primarily focus on parcel deliveries to private customers. Hence, this study aims to develop an agent-based model that explicitly represents the in- and out-going parcel volumes of companies in urban areas delivered by CEPSPs. An approach based on Open Data and self-conducted expert interviews with CEPSPs is developed. First, OpenStreetMap data is used to geographically represent companies with the corresponding sector assignment within a study area in Karlsruhe, Germany. Second, a concept for modeling the weekly in- and out-going parcel volume for each company in the study area is developed using literature-based data. The approach is integrated into the existing agent-based framework *logiTopp* considering all relevant CEPSPs of the respective area. The application shows that modeling CEP-based transportation volumes of companies based on Open Data is possible and leads to reasonable results even though restrictions apply to the granularity of the used data. However, potential is seen in generating a well-funded empirical database of companies' in- and out-going parcel demand structures to improve the model further.

Keywords: urban freight, B2B parcel volumes, agent-based, open data, travel demand

INTRODUCTION

In recent years, there has been substantial growth in e-commerce in Germany, Europe, and the world. The COVID-19 pandemic has accelerated this effect and has led to a further increase in e-commerce activities. While ordering goods via online retailers is a fast and convenient way of shopping for private customers and companies, rising parcel volumes are placing a growing strain on road infrastructure. Courier, express, and parcel service providers (CEPSP) need more vehicles to deliver or pick up parcels to or from their customers. Especially in urban areas, existing space conflicts are intensified, while emissions are steadily rising. Further growth in parcel shipments is expected in the future. Along with the trend of continuing urbanization, it can be expected that the transportation effects caused by courier, express, and parcel (CEP) shipments will intensify.

To be able to investigate the transportation effects of urban goods movement in more detail and to evaluate possible alternative concepts when parcel volumes further increase, freight demand models are used in transportation planning. Both aggregated and disaggregated approaches can be applied. Disaggregated models, in particular, can be used to simulate the relationships between the demand for goods shipments and the resulting transportation effects in more detail, even up to the level of individual parcels. However, a prerequisite for reliable modeling is the availability of accurate input data.

Nevertheless, the availability of sufficient data currently poses challenges to researchers in modeling freight transport in detail, particularly freight volumes. Individuals can be interviewed relatively easily about their e-commerce behavior (e.g., number of orders) via empirical surveys. The literature, therefore, provides multiple examples in which CEP shipments for private customers (B2C[1]) are modeled, especially to analyze the effects on the last mile issue. However, it is difficult to ascertain which goods are delivered to companies by CEPSPs and in what quantities. Microscopic open data sources do not exist either, so the explicit modeling of deliveries to companies by CEPSPs (B2B[2]) on the last mile is insufficiently researched. In Germany, for example, the B2B sector alone is responsible for more than one-third of the total parcel volume and thus accounts for a non-negligible share of the traffic (*1*).

Therefore, this work aims to develop an agent-based model that explicitly represents the in- and out-going parcel quantities of companies in urban areas delivered by CEPSPs. As a modeling environment, we resort to the agent-based travel demand model (TDM) *mobiTopp*, which has already been complemented by the extension *logiTopp* that represents the parcel demand of private customers. We first develop a method that microscopically maps the companies in a study area based on Open Data. Demand is modeled over a week by the number of parcels, which can be flexibly grouped into different shipments. In addition to demand, the volume of parcels generated by companies is also modeled, providing the possibility to consider transportation effects on the last and the first mile. Delivered parcels are also regarded as consumed quantities, picked-up parcels as produced in the following.

This paper is organized as follows. First, we provide an overview of relevant literature on freight demand models. Second, we introduce the agent-based TDM *mobiTopp* and its extension *logiTopp*, which we applied in our approach. Third, the data sources used for our approach and the data preparation are described. Fourth, we explain the model concept and its implementation in *logiTopp*. Finally, we show the model's results, discuss them and give a conclusion.

[1]Business-to-Consumer
[2]Business-to-Business

LITERATURE REVIEW

One of the first agent-based urban freight models is the GoodTrip model. Like most of the following modeling frameworks, it integrates logistical decisions of various actors (policy maker, shipper, receiver, transporter) along the distribution chain to estimate freight transport and its environmental impact (*2*). Limited to the inbound goods flows of the food retail sector and bookstores, the model calculates the consumption volume per goods type in m^3 based on consumer expenditure, which is determined by the spatial distribution of activities and their respective market share. However, the authors do not further define shipment characteristics.

Implementing operation research and logistic sciences concepts in freight modeling, Wisetjindawat et al. (*3*) proposed an agent-based modeling framework for urban freight movement in the Tokyo Metropolitan Area, including commodity production. Each company's production and consumption quantities are estimated according to their commodity type using a regression model on the company's characteristics. These include the location, floor area, and the number of employees, which are generated using the Monte Carlo simulation technique and aggregated distributions of establishments in the study area.

Years later, de Bok and Tavasszy (*4*) developed an agent-based simulation framework for urban freight transport patterns (MASS-GT) based on extensive descriptive statistics of 30,000 observed B2B freight transport operations in Rotterdam, Netherlands. The generation of freight is estimated in yearly volumes (weight) for freight produced and consumed using SMILE+, developed by Bovenkerk (*5*) and Tavasszy (*6*). This module is based on an input-output framework and translates economic scenarios in regional freight production and consumption forecasts for domestic, import, and export freight flows. According to the observed shipment size distribution and companies' characteristics such as size, location, and industry type, the shipments in terms of numbers and weight are simulated between producing and consuming companies.

Schroeder et al. (*7*) extended the passenger transportation modeling framework MATSim by introducing carrier agents into the simulation. A vehicle fleet and depot are assigned to each carrier to execute B2B freight movements from shippers to receivers. Static contracts define the quantity and type of goods to be transported, the delivery and pick-up time windows, and origin and destination locations. Bean and Joubert (*8*) further integrated receivers as autonomous counterparts to carriers enabling dynamic interaction modeling between receivers and carriers. This novelty allows reordering behavior of receiver agents, who could adjust their delivery frequency, - unloading times, and - time window duration throughout the simulation. Also using MATSim as a modeling framework, Llorca and Moeckel (*9*) convert aggregate commodity flow data into long-distance freight flows performed by trucks in the first step. Subsequently, the total production and consumption of parcels are obtained from the resulting number of trips.

Following a holistic logistics approach integrated into the MATSim equivalent SimMobility, SimMobility Freight acts in a multi-scaled simulation framework focusing on B2B freight transportation (*10*). Noticeably, freight generation includes identifying companies as active or non-active regarding freight activities and various combination options for produced and consumed commodity types through a multinomial logit model. Depending on the type of commodity and the separately modeled shipping frequency, either full-truckloads, less-than-truckloads, or parcel shipping is determined.

Intending to analyze the energy consumption of private shopping trips and delivery tours on an urban and national scale, Stinson et al. (*11*) incorporate e-commerce demand and supply into the existing agent-based modeling framework POLARIS. The attributes for the estimation of freight demand are comparable to these mentioned in (*3*).

Considering data collection methods, freight and its generation can be measured by many metrics such as induced vehicle trips, value and number of commodities transported, and the number of stops and deliveries made. According to Holguín-Veras et al. (*12*), freight generation and freight trip generation must be considered differently in data collection. While the former depicts the amount of demand usually measured in tons, the latter represents the number of transportation trips resulting from logistical decisions such as shipment size and frequencies.

Generally, the data stems from national surveys collected by the federal government. This kind of data is usually aggregated and does not distinguish between urban and non-urban freight, though it can be disaggregated under certain conditions (*13*). However, data granularity and accuracy depend on the budget invested in such campaigns (*14*). Usually, the better accessible trip flow data from general surveys is taken in combination with establishment distributions to conclude the distribution of demand (*15*). However, demand is estimated based on certain vehicle types. General surveys aim to collect global information on generation and transport flow variables. In contrast, stakeholder-specific surveys refer to a given category of stakeholders such as companies and transport service providers. Focusing on vehicle usage and driver practices, vehicle-specific surveys contain trip diaries, GPS data, vehicle observations, and driver interviews to provide insights. Furthermore, descriptive area-specific surveys use roadside interviews, traffic counts, and parking observations to complement the information generated with the techniques mentioned above (*16*).

As demonstrated above, existing freight generation modules for companies in agent-based models rely mainly on aggregated freight flow patterns for various commodity types. The explicit consideration of disaggregated demand in terms of parcels that are usually transported by multiple CEPSPs is not given, and hence, simulating operational effects of consolidating parcels to shipments are disregarded. This arises the necessity for a new modeling approach explicitly addressing CEP-based parcel transportation caused by companies. Moreover, existing data sources for freight generation are mainly aggregated, costly to access, and typically vehicle based instead of demand-driven. Consequently, the approach developed in this study relies on independent Open Data making the approach easily transferable, and applicable regardless of the budget.

AGENT-BASED TRAVEL AND FREIGHT DEMAND MODELING FRAMEWORK

For our work, we use and extend the agent-based travel demand modeling framework *mobiTopp* (*17, 18*), which is available as an open source project on GitHub (*19*). *mobiTopp* consists of two separate modules: a long- and a short-term-module. Since we only extend the long-term module in this paper, we only give a short overview of the short-term module.

The long-term module generates a synthetic population for the study area. The population consists of households and their individual members; the agents of the simulation. They are drawn from a population-pool provided by a national household travel survey. In this way, every person and household has properties of a reported real-world entity. *mobiTopp* uses the optimization algorithm 'Iterative Proportional Update' (IPU) to match the households drawn from the population pool with general socio-demographic distributions like age x gender or household size distributions. After generating the synthetic population, each agent is assigned a subset of available mobility tools, including driver's license, public transit pass and memberships to mobility services like car or bike sharing. Next, each agent is assigned an activity-schedule, which describes an agent's sequence of activities over the week each with a start date and duration (*20*). The short-term module simulates each agent's planned activities and the intermediate trips as well as the accompanying short-term decisions of destination- and mode-choice. All agents are simulated simultaneously.

mobiTopp was extended by the logistics module called *logiTopp* (*21, 22*) which integrates CEPSPs, delivery agents and last-mile parcel deliveries to private customers into *mobiTopp*. *logiTopp* is also available as an open-source project on GitHub (*23*).

logiTopp generates the private parcel demand in the study area for one week. This parcel demand is determined after the population is generated by the long-term process described above. To estimate the demand, a parcel demand model is applied to each potential recipient (i.e., every person in the population). First, the number of parcels expected by the recipient over the course of one week is determined. After that, the specific properties of each expected parcel can be determined, including the delivery type (home, work, parcel-locker), the planned delivery day, the CEPSP and the delivery base (DB, also called depot). CEPSP and DB are selected based on market share, while the delivery day is drawn from an even distribution excluding Sunday (since there are no deliveries on Sundays in Germany). The number of parcels and the delivery type are determined by means of discrete choice models (the models in (*21*) have been improved upon since), taking the socio-demographics of the recipient into account.

In the short-term module, the last-mile deliveries of these generated parcels are simulated. Due to the agent-based nature of the framework, *logiTopp* allows for interactions between delivery- and private agents (*22*). Also, each recipient can receive parcels from multiple different CEPSPs. Again, we omit further details of the delivery simulation in the short-term module as this paper focuses on the parcel demand generation of companies which takes place at the end of the long-term module.

DATA

The open available data for freight demand modeling, especially CEPSP freight generation, is limited. Consequently, we had to collect data from multiple sources. On the one hand, we collected data on the characteristics of CEP-based deliveries from the demand and supply perspective based on literature and self-conducted expert interviews. Moreover, to microscopically map the companies in our study area, we collected and prepared data from OpenStreetMap (OSM) (*24*).

Literature-Based Data Sources

The lately booming of e-commerce directed the focus of research towards precisely examining CEPSPs' parcel deliveries in urban areas. While studies on parcel consumption derived from private households received a fair amount of attention (e.g., (*21*)), business consumption mainly got neglected. However, it accounts for over one third of the CEP market, and therefore, Thaller et al. (*25*) analyzed the consumption and production behavior of parcels of commercial recipients and their use of CEPSPs. They rolled out an empirical survey in the area of Berlin, addressing business establishments (including retail, gastronomy, and services) and administration facilities (e.g., research institutions, associations, and public administration units). By conducting a questionnaire-based survey of company owners, the authors collected primary data from 431 companies meeting the requirements of minimal representativeness. Figure 1 shows the yearly amount of delivered and shipped parcels differentiated by sectors, as one of the main results of the study.

Another data source in literature originates from the 'Bundesverband Paket und Expresslogistik' (BIEK) (*1*), where leading CEPSPs in Germany are organized, such as DPD and UPS. Since 2004, BIEK has been publishing an annual CEP study, regarded throughout Germany as the most important series of publications providing a comprehensive description of the CEP market

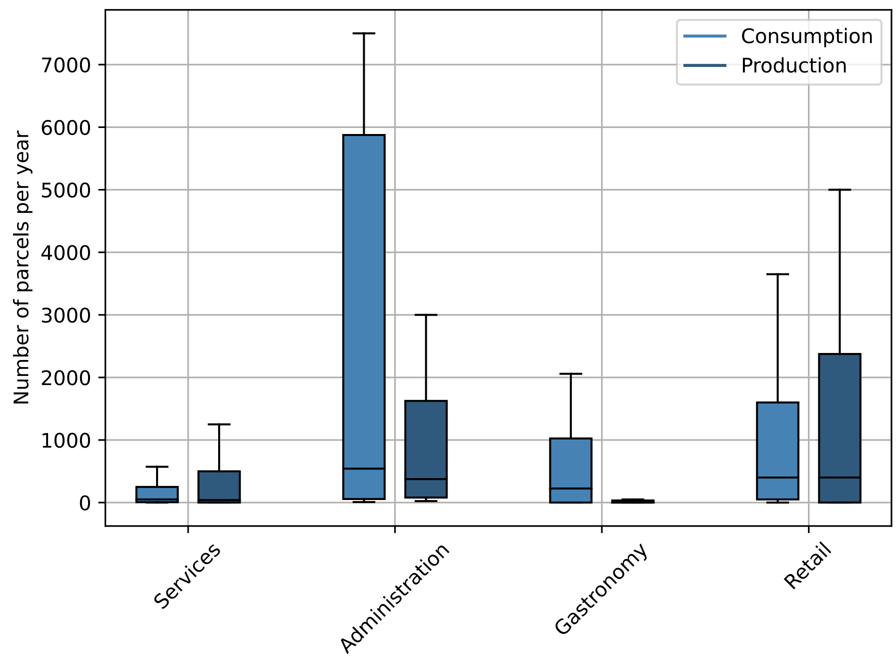

FIGURE 1 Number of consumed and produced parcels per year and sector based on Thaller et al. (*25*)

and is recognized by politics, businesses, the media, and academia. Essential measures, such as the total amount of parcel shipments in Germany (4.51 billion/year) and distribution statistics for B2B-, B2C- and C2C parcels (59% B2C; 37% B2B; 4% C2C) are retrieved from the CEP study 2022, which are used in the later model.

Expert Interviews

To understand CEPSPs' operations and infrastructure, interviews with experts from the leading CEPSPs DHL, Hermes, UPS, and FedEx, usually experienced local branch managers, were conducted. The interviews took, on average, around 1 ½ h and were structured into chapters reflecting the CEPSPs' operations starting from parcel receipt to delivery. In that context, several delivery characteristics, qualitative process descriptions, and customer requirements, such as weekly consumption distribution for receipt and delivery, were consulted. While we have retrieved comprehensive information about parcel consumption and delivery service, we want to share only relevant information in this paper. There is strong agreement across all experts that private households tend to online shopping on weekends. This results in a peak in business parcel receipts on Mondays, representing around a quarter of the total volume. In the following days, the daily volume asymptotically adjusts to the weekly average while dropping firmly on Saturdays due to the weekend rest of a large part of companies. Nevertheless, an observed variation between the operating hours among the service providers, as well as other differences in operational approaches can still lead to slight differences in the weekly distribution. Due to unpredictable internal and external events such as vehicle damage, suddenly occurring driver shortages, poor weather conditions, or abnormally high parcel handover times (e.g., elevator reliance) during the day, the delivery quantities can deviate from the planned.

OpenStreetMap Data

Microscopic modeling of companies' in-going and out-going parcel volumes delivered by CEP-SPs requires information about the number, type, and location of all companies within the study area. To achieve transferability of our methodology to other regions, we gained this information from the Open Data source OpenStreetMap (*24*). In OSM, objects are geographically mapped and supplemented with varying feature variables of the object itself. Although OSM data is not expected to provide information about all companies within the study area, we assumed sufficient data quality because, e.g., in Germany and the U.S., OSM data is provided by public authorities and not only by volunteers. Moreover, other studies such as Ziemke et al. (*26*) have already shown the successful utilization of OSM data, for example, for the synthetic passenger transport demand generation.

For our study, we adapt the procedure introduced by Klinkhardt et al. (*27*). They extracted all points of interest of a designated area which can be easily adapted to company buildings. First, we analyze which feature variable of OSM objects is relevant to identify only company-related objects. On the one hand, we select relevant feature variables such as "shop" or "office," which can directly be classified as a company. On the other hand, we also select feature variables that allow us to classify an object as a company only after carefully reviewing further tag information of an object. As an example, the feature variable "amenity" can be tagged as "car_parking" or "bar", whereas only the latter is a company-related object. Based on an OSM export of our study area, we use the open source application 'Osmosis' to extract relevant OSM objects by applying the previously defined feature variables as filters. The output list is cleaned in an extensive and partly manual data review process by eliminating objects that cannot be recognized as a company based on feature variables' values. In total, 7,037 objects are extracted and identified as companies in the study area of Karlsruhe, Germany.

Next, the final 'Osmosis' output file is imported to ArcGIS. Two crucial data processing steps are carried out. First, all companies are allocated to the corresponding travel analysis zone (TAZ) based on the extracted geographical coordinates. Second, based on the polygon surfaces in which a company is located, the floor area of a company in square meters is derived based on the method described in Klinkhardt et al. (*27*). If the information about a company's number of floors is available in the OSM export, it is also regarded. However, it is not possible to determine a floor area for all companies, so missing values are imputed.

As a final pre-processing step, the extracted companies are assigned to branch categories based on the standardized NACE classification (*28*) to account for differences between company branches when modeling demand. Which company is assigned to which branch is decided in a schematized way based on the values of feature variables. Case-by-case checks are carried out if the assignment is unclear, e.g., based on the company name. In addition, we assign each branch category to a higher-level sector based on structural similarities of the branches. Table 1 shows the distribution of extracted companies from OSM by branch and sector category and compares it to the actual distribution of companies in the city of Karlsruhe, which was gained from the official registry (*29*).

It becomes clear that for some branches or sectors, the distribution of the OSM extracted companies matches the actual distribution quite closely. This applies especially to companies in *gastronomy*, *retail*, and *leisure* sectors, where our pre-processed OSM export already covers 82% to 94% of actual companies. Greater deviations are observable for the sectors *industry* and *service*. However, these results are in line with literature. Briem et al. (*30*) found that objects open to the

TABLE 1 Branch and sector based comparison of OSM companies with statistical data of official registry (*29*)

Sector	Branch	Branch based		Sector based	
		Statistics	OSM	Statistics	OSM
Administration	Education	403	515	403	686
	Public Administration and Defense; etc.	0	171		
Gastronomy	Accommodation and Food Service Activities	1,142	1,070	1,142	1,070
Industry	Manufacturing	664	571	2,266	741
	Transportation and Storage	441	92		
	Water Supply; etc.	26	23		
	Construction	1,049	28		
	Electricity, Gas Supply; etc.	86	27		
Leisure	Arts, Entertainment, and Recreation	504	441	504	441
Retail	Wholesale and Retail Trade; etc.	2,376	1,944	2,376	1,944
Service	Administrative and Support Service Activities	979	75	8,392	2,155
	Financial and Insurance Activities	383	204		
	Human Health and Social Work Activities	1,087	495		
	Information and Communication	1,166	193		
	Other Service Activities	1,114	506		
	Professional, Scientific, and Techn. Activities	2,970	655		
	Real Estate Activities	693	27		
		15,083	7,037	15,083	7,037

public are better maintained in OSM than, for example, small offices, which are primarily classified as *service* objects in our approach. *Administration* companies operated by the government are not considered in the official registry, which explains why our procedure identifies more objects than reported in the actual statistics. In summary, we conclude that our extraction and classification procedure of OSM objects brings up plausible results.

However, for the later model, the OSM objects must be aligned with the overall number and distribution of officially registered companies. For this purpose, we divide the extracted companies from OSM into sector-specific sub-groups. Each sub-group is set as a distinct base. Then, the missing number of objects are generated, which is the difference between the number of objects in the sub-group and the reported number of companies from the official registry. Consequently, we apply a sampling technique in the data software *R*. With this, the defined number of objects is drawn from the base objects, thereby accounting for structural patterns from the sector-specific base group. As an example, for the *retail* sector 432 objects were drawn from 1,944 OSM objects. The *administration* sector poses an exception. Since no number of companies is reported for 'Public administration', we adopted the number of OSM objects after manually checking the plausibility of the extracted number of objects. In total, we generated a synthetic data set comprising 15,366 company objects.

MODELING APPROACH

All the data described in the previous section has been used to develop the concept for microscopically modeling the parcel shipments of companies in urban areas that are delivered or picked up by CEPSPs. First, this section explains the modeling concept in more detail. Second, the steps necessary to implement the concept in the existing modeling framework *logiTopp* are specified.

Concept

In contrast to most existing studies, we decided to model parcels as the smallest logistical unit instead of shipments. However, parcels are consolidated into shipments guaranteeing a more flexible but realistic model enabling several parcels of one company to be split into several shipments. Each shipment composes a minimum of one parcel and a maximum of the total parcel quantity of a company. Moreover, the modeling horizon is set to one week. Even though the results from the expert interviews in Section 5.2 have shown that there is hardly any variation over the weekdays in the delivery of parcels to companies, the variation in the pick-up of parcels is tremendous, making a modeling horizon greater than one day, as usually in other models, necessary.

Our model for representing CEPSP-based parcel shipments caused by companies aims at a holistic approach, illustrated in Figure 2. On the transport demand side, the model determines the parcel quantities companies require to transport by CEPSPs. We explicitly distinguish between the delivery and pick-up process. Hence, we model the number of consumed and produced parcels per company, which is again the volume of parcels delivered and picked up by CEPSPs. Especially to account for the transportation-related effects of e-commerce activities, it is essential to consider the produced parcels as nearly each e-commerce order 'produces' parcels at a company, i.e., in the retail sector. Moreover, varying opening hours of companies are considered as these determine the time in which parcels can be delivered or picked up. On the transport supply side, the concept considers the microscopic representation of CEPSPs. To reflect the effects of different organizations, several CEPSPs are represented, whereas parcels are distributed based on the market share of each CEPSP. With this, since the model also allows for consideration of CEPSP-specific vehicle fleets, the effects of varying transport capacities between different CEPSPs' can also be reflected. Parallelly to the opening hours of companies, for each CEPSP specific operating hours are assigned.

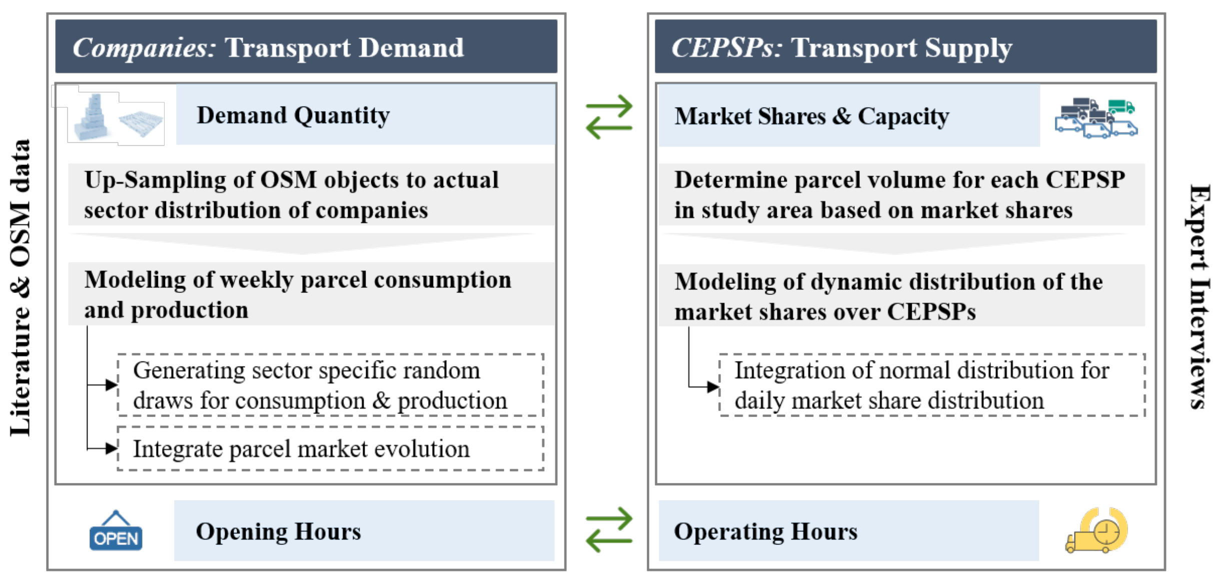

FIGURE 2 Concept of parcel generation model for companies

The modeling of the transport demand side in Figure 2 is mainly built on the literature-based data and the structural data of OSM objects. First, the companies extracted from OSM are up-sampled to the actual distribution stated by the official registry as described in the previous section. Second, each company's weekly parcel consumption and production volume is modeled. Thereby, we apply the results gained in the research of Thaller et al. (*25*), which was carried out in a German urban area and is suitable for our study. However, the data is only available on the aggregated level of box plots, as shown in Figure 1. Nevertheless, boxplots provide information on the distribution of a variable, namely the minimum, lower quartile, median, upper quartile, and maximum values. Consequently, we use these figures to replicate the distribution of produced and consumed parcels and apply the distribution to the list of companies in our study area.

For the sectors *administration*, *gastronomy*, *retail*, and *service*, the key figures of the box-plots can be adopted directly. *Leisure* and *industry* are not included in the study by Thaller et al. (*25*). As these are composed of elements of all other sectors, we take the mean values of other sectors considering structural differences between consumed and produced quantities. The appropriate number of uniformly distributed random draws between the corresponding key values is determined for each sector. For example, for all 504 *leisure* objects, 126 random draws are generated between the minimum and lower quartile, lower quartile and median, and so on. This procedure is carried out for both production and consumption quantities of parcels, and resulting quantity vectors are paired sector-wise. Next, the produced and consumed parcels are broken down into weekly quantities as they are reported annually in Thaller et al. (*25*). Moreover, as the study was carried out in 2018, the parcel quantities are adjusted according to the annual evolution of the CEP market (c.f. (*1*)) to model quantities of 2021. Finally, production and consumption quantities are matched sector-wise with companies based on the corresponding floor area. Companies with smaller floor areas are more likely to match smaller parcel quantities. In addition, each company gets assigned the same opening hours by default.

The transport supply offered by CEPSPs is modeled based on expert interviews and further research. Locations of all DBs of all CEPSPs within the study area are collected and considered in the model. Furthermore, market shares of the different CEPSPs are integrated into the model controlling which CEPSP transports which quantity of parcels. In the expert interviews, it became clear that the daily transport quantity usually varies by about 11% more or less than the planned quantities. Consequently, transportation quantities based on the market share are modeled dynamically, considering a normal distribution with the market share as the mean and the variation index as the standard deviation. Additionally, each CEPSP is represented with its operating hours.

The presented concept is embedded in a greater framework, in which mid- and short-term processes such as vehicle allocation and delivery or pick-up tours are modeled. However, this is not the focus of the study on hand and hence, not presented.

Integrating companies as recipients and shippers into logiTopp

We integrate our approach and the models described above into the agent-based travel demand modeling framework *mobiTopp* i.e., its logistics extension *logiTopp*. In its current state, *logiTopp* only supports private persons as recipients and all parcels are delivered by CEPSPs.

To integrate our approach, we extend the framework by introducing companies as a third type of agent into the delivery process. Each company in the survey area is modeled explicitly as a unique agent with a name, branch (according to the NACE classification (*28*)), sector, building type, floor area in square meters, opening hours per day as well as a coordinate and the corre-

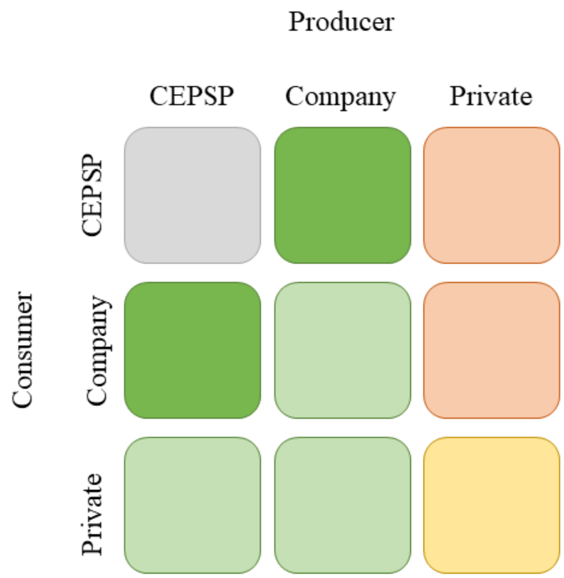

FIGURE 3 Possible producer-consumer relations: dark-green - supported by *logiTopp* and investigated in this paper, light-green: supported but not investigated, yellow - implicitly modeled by *mobiTopp*, red - not yet supported by *logiTopp*, gray - irrelevant

sponding TAZ. Like private persons, companies can act as recipients of parcels. However, they are also able to produce goods and, therefore new parcels. In this way, companies both produce and consume parcels, making the previously trivial delivery chain (merely CEPSP-to-Private as described in Section 4) more complex. Therefore, we abstract from the three types of agents and introduce two roles: producer and consumer. In this abstraction, all three types of agents: CEPSP, Company and Private (person) can fulfill both roles; however not all producer-consumer relations are supported by *mobiTopp* or even relevant.

The matrix shown in Figure 3 presents the different producer-consumer relations and how they are integrated in *logiTopp*: CEPSPs deliver parcels to companies and private persons (but not mutually); they act as the producer since they are the parcels' entry points into the survey area and insert these parcels into the simulated system.

Companies can also produce parcels and send them outside the survey area: in this case, CEPSPs are the exit points for these parcels leaving the simulated area. Hence, CEPSPs are the consumers in this context as they represent an outside demand and take the parcels out of the simulated system. Furthermore, companies can deliver their produced parcels directly to the company or private person within the study area. The software supports both relations; however, they are not yet fully modeled with data of the study area.

Private persons could also act as producers themselves, e.g., returning parcels to their producer (CEPSP or company). These two relations are not yet supported by *logiTopp*. Finally, private persons can deliver parcels to other private persons, e.g., when privately selling things on platforms like eBay. This relation can also not be modeled explicitly in *logiTopp*; however, the resulting travel demand is implicitly modeled by *mobiTopp*'s 'private-business' activities.

In this paper, we investigate the relations CEPSP-Company and Company-CEPSP (marked dark-green in Figure 3). Therefore we provide a parcel-consumption model, and a parcel-production model for companies. The production and consumption quantity is determined according to the data and models described in Sections 5 and 6.1. For each parcel to be received by a company agent, three properties are determined: the planned day of arrival, as well as a CEPSP and one of its DBs. The arrival day is drawn from the production/consumption week distribution obtained from the expert interviews as described in Section 5.2. Both the CEPSP and one of its DBs are determined based on their market share. If the company at hand produces a parcel, the DB serves as a consumer, otherwise it serves as a producer.

RESULTS AND DISCUSSION

The previously described model has been applied to the study area of Karlsruhe, Germany. The consumed and produced parcel quantities have been simulated for all 15,366 companies and distributed over a week. The simulated area is serviced by the six major CEPSPs, namely DHL (48%), Hermes (16%), UPS (12%), DPD (10%), GLS (7%), and FedEx (6%), which are integrated into the model with the location of their DBs. The figures in brackets represent the corresponding market shares based on parcel volumes according to Pitney Bowes (*31*).

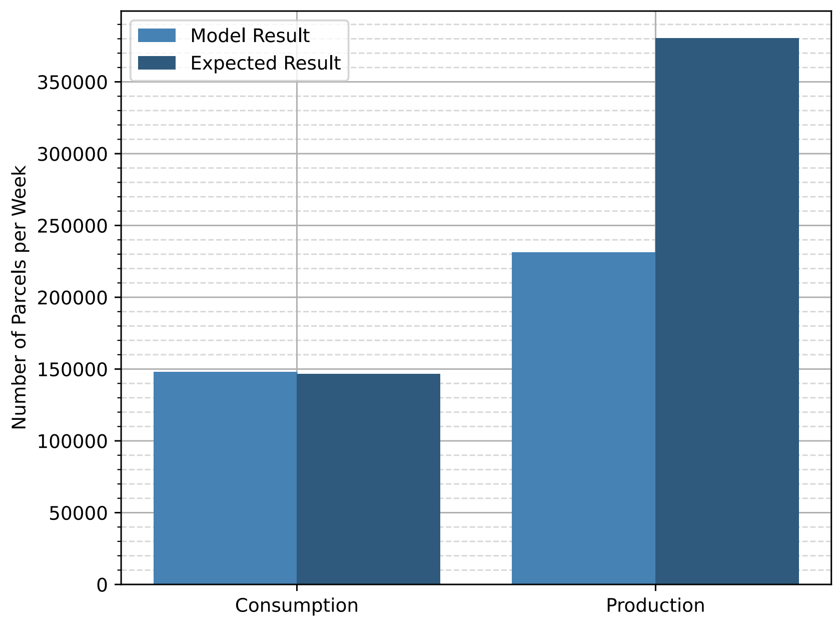

FIGURE 4 Comparison of total produced and consumed parcel quantities between model and expected results; expected results based on (*1*)

In total, the model generates 147,833 parcels delivered to companies and 231,298 parcels that need to be picked up at the company sites. In Figure 4, these numbers are illustrated and compared with the expected quantity of parcels based on the study of BIEK (*1*). As the latter only provides the respective quantities for the overall German market, the numbers were down-scaled based on the share of companies within Karlsruhe out of the total number of companies in Germany. We can see that our model meets the consumed parcel quantity very close. There is only a slight fluctuation of about 1,000 parcels. Since the expected quantity is based on a down-scaling and not a directly reported or surveyed figure in the study area, it is also subject to uncertainty, and thus, deviations were expected. However, this cannot explain the great gap of about 150,000 parcels between the expected and modeled volume of produced parcels. This is because the quantity reported in the BIEK study also includes the volume of parcels generated by large e-commerce retailers such as Amazon or Zalando. Amazon alone is responsible for about 10% of all parcels produced in Germany (*32*). Our study area is limited to the urban area of Karlsruhe, where no such companies are represented, which explains a significant underestimation of the expected volume reported here. This effect does not affect consumption as CEPSPs only scarcely deliver to e-commerce retailers.

Figure 5 illustrates the geographical distribution of the total consumption volume of parcels per TAZ. The stronger a TAZ is shaded, the more parcels are consumed. The map shows overall reasonable results. Two darkly shaded areas are observable outside the city center, one in the west and one in the east of the study area. Both areas represent industrial zones in Karlsruhe with a high density of companies. However, darker shaded TAZs are also visible in the city center, especially on the western side. Many retail shops and popular malls are located in this area, again causing a comparatively high number of companies. White shaded TAZs represent mostly recreational areas, where no or only few companies are located, and hence, a low parcel consumption is reasonable.

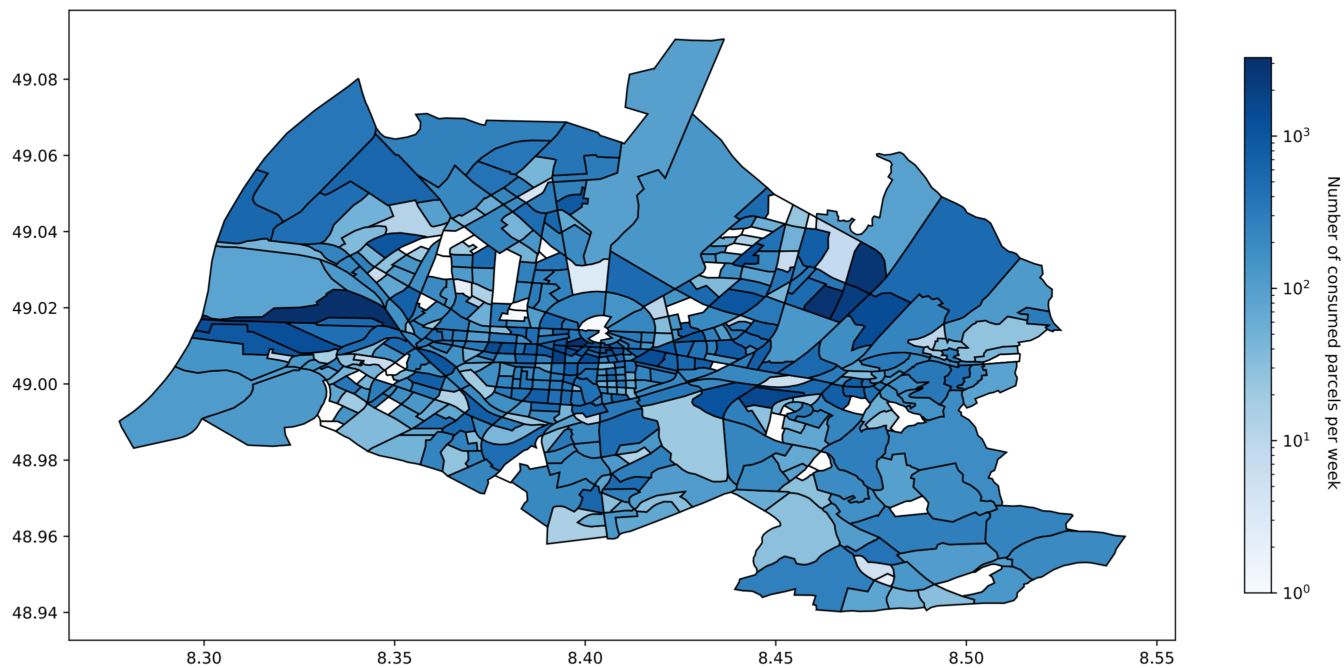

FIGURE 5 Geographical distribution of consumed parcels per TAZ

Moreover, consumed and produced parcels are modeled differently between sectors. Therefore, Figure 6 shows the sector-wise boxplots for modeled parcel consumption and production quantities and directly compares it with the input distribution based on Thaller et al. (*25*) as presented in Figure 1. To achieve comparability, data from Thaller et al. (*25*) in Figure 6 was scaled to weekly quantities and adjusted to the CEP market evolution from 2018 to 2021 according to the modeled parcel volumes. As the sectors *industry* and *leisure* were not considered in the study by Thaller et al. (*25*), no comparison is possible. According to the other sectors, we can see an overall good fit of the sector-specific distribution of modeled and expected parcel quantities for consumption and production. This indicates that the suggested procedure in Section 6.1 is a suitable approximation of actual parcel quantities. However, this comparison only shows the distribution based on minimum, lower quartile, median, upper quartile, and maximum. It is not yet clear if we also meet the distribution between those key figures. More detailed data would be necessary to compare these more detailed distributions.

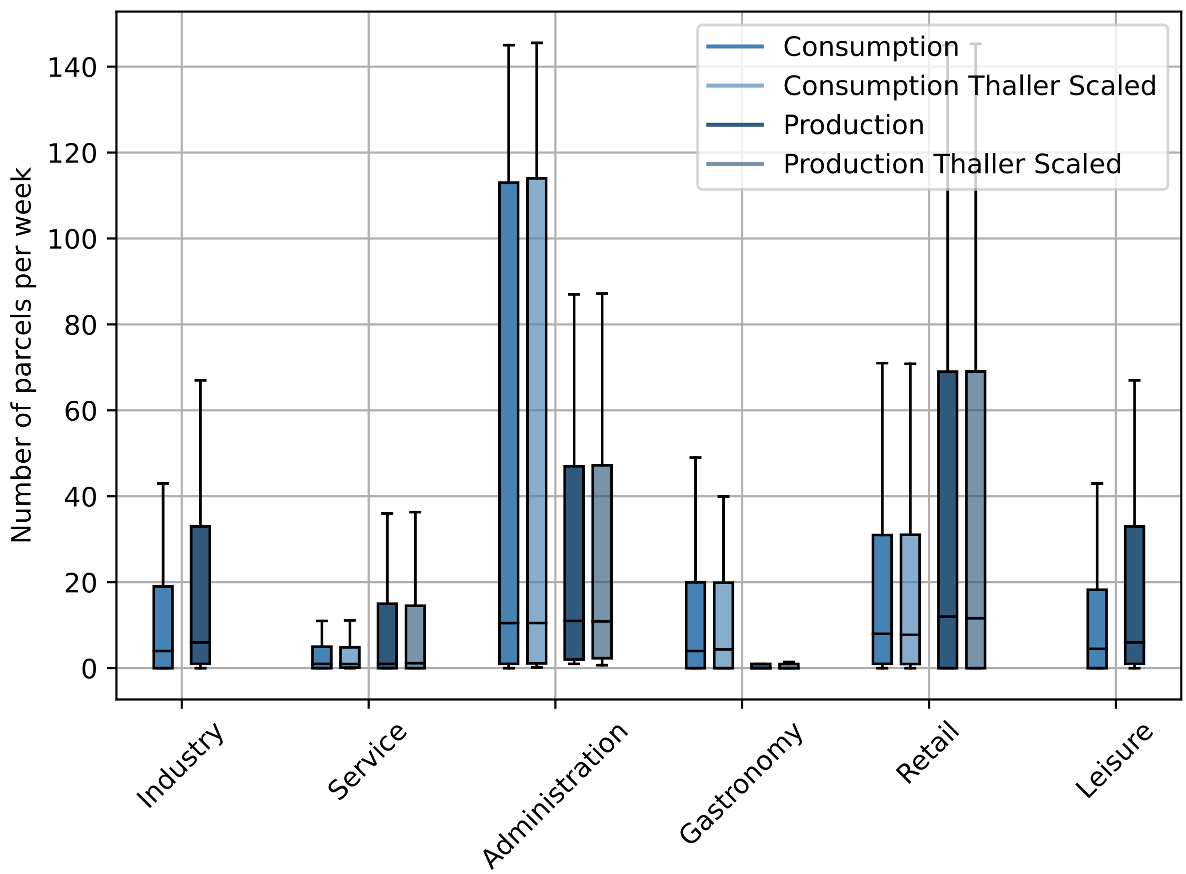

FIGURE 6 Distribution of sector-specific produced and consumed parcel quantities; compared with results of Thaller et al. (*25*)

The model also simulates the distribution of parcel quantities over the week and allocates the quantities to available CEPSP. The results are shown in Figure 7, where on the left side, delivered parcels (based on companies' consumption) and on the right side, picked-up parcels (based on companies' production) are illustrated. We can see that companies receive parcels uniformly distributed over the week, which is reasonable as companies also show a balanced shopping behavior. In contrast, parcel pick-ups follow a different distribution with a peak on Monday that asymptotically drops until Friday. The distribution was calibrated based on findings gained in the expert interviews (c.f. Section 5.2) and represents the peak in private e-commerce orders on weekends that companies have to ship starting from Monday. In both distributions, Saturday forms an exception. To this day, many companies, especially in the service sector (e.g., offices), are closed, which makes the decrease in parcel consumption and production reasonable. Both diagrams also show which CEPSP delivers and picks up which parcel quantities. The distribution reflects the market share as desired. However, a slight difference between days is observable, representing variations due to operational reasons of a CEPSP as presented in Section 5.2.

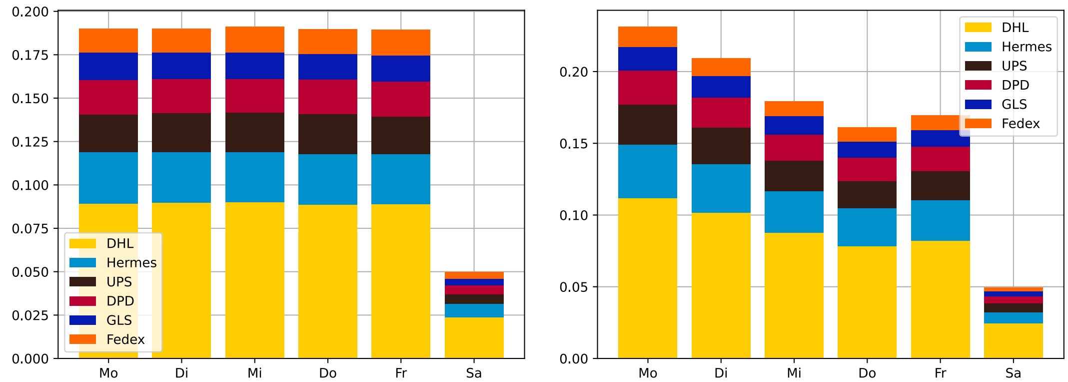

FIGURE 7 Distribution of consumed (left) and produced (right) parcels over the week, displayed by CEPSPs

CONCLUSION

Although B2B deliveries account for a large share of the total parcel volume in the CEP market, they are often not explicitly considered in existing freight demand models, especially in agent-based models. One reason is the lack of sufficient data, which is a crucial problem in freight demand modeling. However, in the present study, we could develop a feasible and flexible agent-based approach to model the CEP-based transportation volumes of companies in terms of produced and consumed parcels based on Open Data. Even though restrictions apply in the granularity of some data, the presented approach shows overall reasonable results.

First, we showed that using OpenStreetMap is a suitable database to collect information about the location of companies as well as other characteristics. However, we recognized that the data quality varies between companies of different sectors. Therefore, it might be helpful to consider secondary data, e.g., from public authorities, to better reflect the locations of companies, especially from the *service* and *industry* sectors. Unfortunately, this data is not easy to obtain. Second, we showed that external, aggregated information on parcel distributions can be used to model parcel volumes on an agent-based level, leading to comprehensive results. However, this requires the availability of data on all relevant objects, which was not the case in our study, e.g., for the sectors *leisure* and *industry*, and had to be circumvented with assumptions whose effects, in turn, cannot be validated.

Besides the technical aspects, our study also emphasized the importance of modeling transportation demand for CEPSPs for more than one day, as usually done in existing models. Especially the out-going parcel volumes of companies are subject to non-negligible fluctuations over one week resulting in different implications for the actual traffic. Moreover, instead of only focusing on the last mile issue, considering parcels produced by companies also allows to address the first mile, which is essential when analyzing the overall effects on the transportation system.

For further research, we will also integrate the mid- and short-term decisions into the *logiTopp* framework. This includes the agent-based consideration of CEPSPs' vehicle fleets and tour-planning characteristics to analyze the transportation effects with varying parcel volumes. Even more important, we see the generation of a well-funded empirical database of companies' in-and out-going parcel demand structures. Although the presented model based on aggregated data showed reasonable results, we still see great potential to improve the representation of the exact parcel distributions. Hereby, we can also contribute to further insights into the parcel structure of sectors where no data was available. We already designed a survey and performed a pre-test. After the field phase, the data is used to update our presented approach.

AUTHOR CONTRIBUTION STATEMENT

The authors confirm contribution to the paper as follows: study conception and design: L. Barthelmes, M.E. Görgülü, J. Kübler, M. Kagerbauer, P. Vortisch; data collection: M.E. Görgülü, L. Barthelmes; analysis and interpretation of results: L. Barthelmes, J. Kübler, M.E. Görgülü; model implementation: J. Kübler; draft manuscript preparation: L. Barthelmes, M.E. Görgülü, J. Kübler. All authors reviewed the results and approved the final version of the manuscript.

ACKNOWLEDGMENTS

The work presented in this paper stems from research within the project "LogIKTram - Logistics Concept and ICT Platform for Urban Rail-based Freight Transport" funded by the Federal Ministry of Economic Affairs and Climate Action of Germany.

REFERENCES

1. BIEK: Bundesverband Paket und Expresslogistik e. V., *KEP-Studie 2022 – Analyse des Marktes in Deutschland*. Berlin, Köln, 2022.
2. Jeroen Boerkamps, Arjan Van Binsbergen, *GoodTrip – A new approach for modelling and evaluation of urban goods distribution*, 1999.
3. Wisetjindawat, W., K. Sano, S. Matsumoto, and P. Raothanachonkun, Microsimulation Model for Modeling Freight Agents Interactions in Urban Freight Movement. *Transportation Research Board 86th Annual MeetingTransportation Research Board*, , No. 07-2224, 2007.
4. de Bok, M. and L. Tavasszy, An empirical agent-based simulation system for urban goods transport (MASS-GT). *Procedia Computer Science*, Vol. 130, 2018, pp. 126–133.
5. Bovenkerk, M., SMILE+: the new and improved Dutch national freight model system. *Proceedings of ETC 2005, Strasbourg, France 18-20 September 2005 - Transport Policy and Operations - Freight and Logistics - Freight Modelling I*, 2005.
6. Tavasszy, L., A DSS for modelling logistic chains in freight transport policy analysis. *International Transactions in Operational Research*, Vol. 5, No. 6, 1998, pp. 447–459.
7. Schroeder, S., M. Zilske, G. Liedtke, and K. Nagel, Towards a multi-agent logistics and commercial transport model: The transport service provider's view. *undefined*, 2012.
8. Bean, W. L. and J. W. Joubert, Modelling receiver logistics behaviour. *Procedia Computer Science*, Vol. 151, 2019, pp. 763–768.
9. Llorca, C. and R. Moeckel, Assesment of the potential of cargo bikes and electrification for last-mile parcel delivery by means of simulation of urban freight flows. *European Transport Research Review*, Vol. 13, No. 1, 2021, pp. 1–14.
10. Sakai, T., A. Romano Alho, B. K. Bhavathrathan, G. D. Chiara, R. Gopalakrishnan, P. Jing, T. Hyodo, L. Cheah, and M. Ben-Akiva, SimMobility Freight: An agent-based urban freight simulator for evaluating logistics solutions. *Transportation Research Part E: Logistics and Transportation Review*, Vol. 141, 2020, p. 102017.

11. Stinson, M., J. Auld, and A. Mohammadian, A large-scale, agent-based simulation of metropolitan freight movements with passenger and freight market interactions. *Procedia Computer Science*, Vol. 170, 2020, pp. 771–778.
12. Holguín-Veras, J., M. Jaller, L. Destro, X. Ban, C. Lawson, and H. S. Levinson, Freight Generation, Freight Trip Generation, and Perils of Using Constant Trip Rates. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2224, No. 1, 2011, pp. 68–81.
13. Allen, J., C. Ambrosini, M. Browne, D. Patier, J.-L. Routhier, and A. Woodburn, Data Collection for Understanding Urban Goods Movement. In *Sustainable Urban Logistics: Concepts, Methods and Information Systems*, Springer, Berlin, Heidelberg, 2014, pp. 71–89.
14. Holguín-Veras, J. and M. Jaller, Comprehensive Freight Demand Data Collection Framework for Large Urban Areas. In *Sustainable Urban Logistics: Concepts, Methods and Information Systems*, Springer, Berlin, Heidelberg, 2014, pp. 91–112.
15. Fekpe, E., M. Alam, T. Foody, and D. Gopalakrishna, Freight Analysis Framework Highway Capacity Analysis: Methodology Report, 2002.
16. Allen, J., M. Browne, and T. Cherrett, Survey Techniques in Urban Freight Transport Studies. *Transport Reviews*, Vol. 32, No. 3, 2012, pp. 287–311.
17. Mallig, N., M. Kagerbauer, and P. Vortisch, mobiTopp – A Modular Agent-based Travel Demand Modelling Framework. *Procedia Computer Science*, Vol. 19, 2013, pp. 854–859.
18. Mallig, N. and P. Vortisch, Modeling travel demand over a period of one week: The mobiTopp model. *arXiv preprint arXiv:1707.05050*, 2017.
19. Kübler, J. and L. Briem, *mobiTopp*. https://github.com/kit-ifv/mobitopp, 2022.
20. Hilgert, T., M. Heilig, M. Kagerbauer, and P. Vortisch, Modeling week activity schedules for travel demand models. *Transportation Research Record*, Vol. 2666, No. 1, 2017, pp. 69–77.
21. Reiffer, A., J. Kübler, L. Briem, M. Kagerbauer, and P. Vortisch, Integrating Urban Last-Mile Package Deliveries into an Agent-Based Travel Demand Model. *Procedia Computer Science*, Vol. 184, 2021, pp. 178–185.
22. Jelle Kübler, Anna Reiffer, Lars Briem, and Peter Vortisch, Integrating Neighbours into an Agent-Based Travel Demand Model to Analyse Success Rates of Parcel Deliveries. *Procedia Computer Science*, Vol. 201, 2022, pp. 181–188.
23. Kübler, J., L. Barthelmes, M. E. Görgülü, and A. Reiffer, *logiTopp*. https://github.com/kit-ifv/logitopp, 2022.
24. OpenStreetMap contributors, *OpenStreetMap.org*, 2022.
25. Thaller, C., L. Papendorf, P. Dabidian, U. Clausen, and G. Liedtke, Delivery and Shipping Behaviour of Commercial Clients of the CEP Service Providers. Springer, Cham, 2019, pp. 205–220.
26. Ziemke, D., I. Kaddoura, and K. Nagel, The MATSim Open Berlin Scenario: A multimodal agent-based transport simulation scenario based on synthetic demand modeling and open data. *Procedia Computer Science*, Vol. 151, 2019, pp. 870–877.
27. Klinkhardt, C., T. Woerle, L. Briem, M. Heilig, M. Kagerbauer, and P. Vortisch, Using OpenStreetMap as a Data Source for Attractiveness in Travel Demand Models. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2675, No. 8, 2021, pp. 294–303.

28. Eurostat, *Statistical Classification of Economic Activities in the European Community (NACE)*, 01.01.2008.
29. Hammer, A., D. Götz, B. Dierßen, and K. Rohweder, *Unternehmens- und Betriebsumfrage 2021*. Karlsruhe, 2021.
30. Briem, L., M. Heilig, C. Klinkhardt, and P. Vortisch, Analyzing OpenStreetMap as data source for travel demand models A case study in Karlsruhe. *Transportation Research Procedia*, Vol. 41, 2019, pp. 104–112.
31. Pitney Bowes, *Parcel shipping index 2021*. Stamford, United States, 2022.
32. Statista, *Amazon - Paketsendungen nach Versandanbieter 2022 | Statista*, 01.08.2022.