Value of Household Availability Profiles for Last-Mile Routing

Maximilian Engelhardt
School of Computing, Communication, and Business, University of Applied Sciences Berlin, Germany
maximilian.engelhardt@htw-berlin.de

Stefan Voigt
WFI - Ingolstadt School of Management, Chair of Logistik & Operations Analytics, KU Ingolstadt, Germany

Stephan Seeck
School of Computing, Communication, and Business, University of Applied Sciences Berlin, Germany

Sebastian Lang
Institute for Engineering of Products and Systems (IEPS), Otto von Guericke University Magdeburg, Germany

Abstract
Failed deliveries are a major challenge in last-mile logistics, increasing costs and emissions and reducing service quality. Household availability prediction offers a promising approach by aligning delivery attempts with the likelihood of successful handover. However, its real-world value remains unclear. This paper evaluates the impact of household availability profiles (APs) using empirical data and optimization-based experiments. APs are derived from the German Time Use Survey and combined with real shipment data from a German logistics service provider. Routing instances are solved using the Vehicle Routing Problem with Availability Profiles (VRPAP) under varying levels of availability information (0%–100%). Results show that increasing AP coverage significantly reduces failed deliveries and improves service performance. At the same time, travel distances remain stable, indicating no increase in routing inefficiency. However, routing costs rise due to time-based effects, particularly as waiting times are allowed and not fully captured in the cost modeling approach.

1. Introduction
In last-mile logistics, failed deliveries are a significant operational challenge, leading to additional costs and emissions, lower service quality, and customer dissatisfaction. Up to 20% of B2C deliveries fail, primarily caused by customers not being home (Lim et al., 2023; van Duin et al., 2016). Considering that the last mile is the most expensive segment of the supply chain, accounting for approximately 50–60% of total supply chain costs (Bundesverband Paket- und Expresslogistik e. V. (BPEX), 2023), underscores the importance of the last mile and the problem of failed deliveries. Existing solutions, such as parcel lockers or time windows, either conflict with customers' preferences for personal handovers (Peppel et al., 2022; Schröder et al., 2018) or increase costs (Boyer et al., 2009). A promising data-driven approach overcomes these obstacles by leveraging customer availability profiles in operational last-mile planning and execution to increase the number of successful home deliveries. Previous research indicates that this new technology reduces last-mile costs by 5%-40% by balancing the trade-off of increasing time- and distance-based tour costs and decreasing failed delivery costs (Escudero-Santana, 2026; Seghezzi and Mangiaracina, 2023; Özarık et al., 2023, 2021; Voigt et al., 2022; Praet and Martens, 2020; Florio et al., 2018; Pan et al., 2017). However, the assumptions in these studies are stylized, and the author's discussions with practitioners revealed concerns about whether reduced failed-delivery costs justify detours in real-world settings, especially where neighbors accept deliveries or safe places are available. Thus, the real-world value of this new promising technology for last-mile operations remains unclear. Therefore, this project aims to investigate the value of customer availability prediction technology in last-mile operations in more realistic settings than previously studied. Based on the described research problem, the main research question is: "How valuable is the customer availability prediction technology for last-mile operations in terms of cost- and service-effectiveness and sustainability by reducing failed deliveries in realistic settings?"

http://dx.doi.org/10.25673/123556

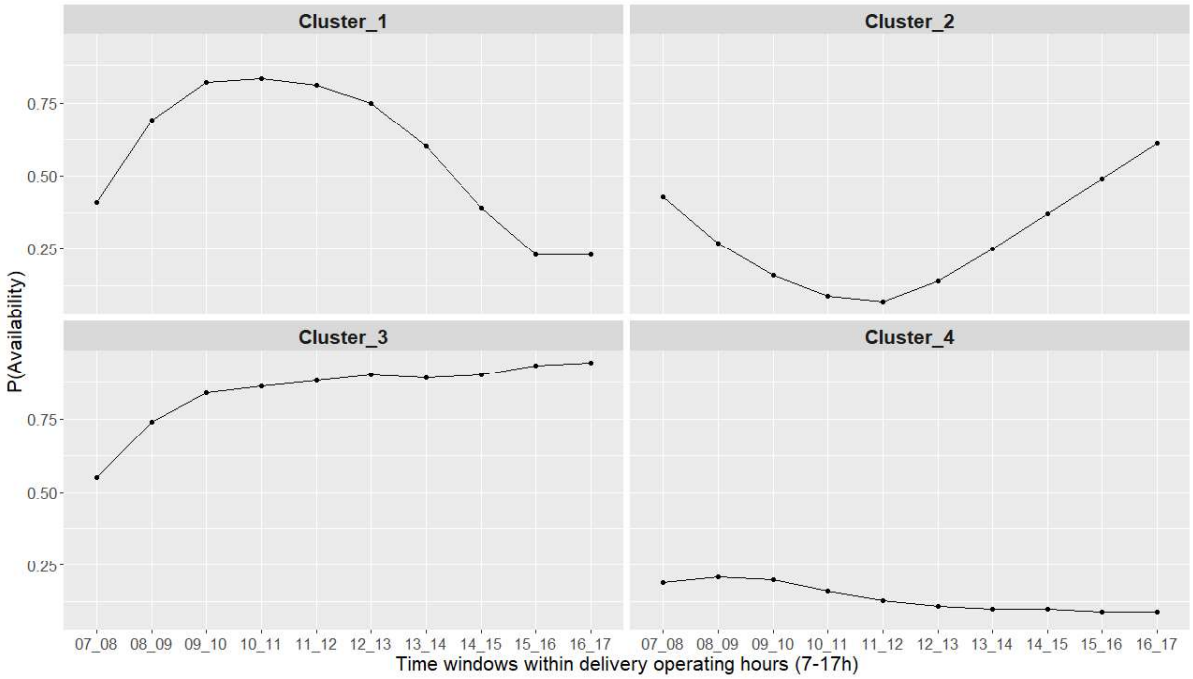

Figure 1: Empirical Household Availability Profiles based on the 2022 German Time Use Survey

As in practical settings, household availability data might not be available for all customers, the objective of this paper, which is an initial result of the project, is to assess how the availability of household availability data impacts last-mile delivery performance in realistic settings. The missing-availability-profile issue represents a primary real-world barrier that was not examined in previous studies but emerged as a key barrier in the author's discussions with practitioners, as customers may not be willing to share availability information for privacy reasons. Therefore, last-mile operations are simulated with and without APs using empirical shipment data and mathematical optimization approaches. This research contributes by investigating the problem experimentally using empirical data. In contrast to the closest existing studies, our study provides new insights into the issue of missing APs and their impact on the operational outcomes. Our study does not contribute to mathematical optimization, as we use the VRPAP model and solution algorithm of Voigt et al. (2022) only as an experimental evaluation instrument.

2. Methods and Experimental Design

To determine the value of customer availability profiles for last-mile routing in realistic settings, the following experimental approach is used. First, we derive customer availability profiles empirically from the 2022 German Time Use Survey (Statistisches Bundesamt, 2024). Second, we prepare empirical data on last-mile shipments from a German logistics service provider and map customer availability profiles to tour stops, which serve as input instances for our experiments. Third, we use the algorithm of Voigt et al. (2022) to solve the vehicle routing problem with availability profiles for our empirically-derived instances.

2.1. Empirical Customer Availability Profiles

To obtain realistic customer availability patterns, we derive empirical availability profiles from the 2022 German Time Use Survey (Statistisches Bundesamt, 2024). This dataset provides detailed time-diary information on individuals' daily activities and locations in 10-minute intervals over a full 24-hour period. Compared to stylized assumptions used in prior research, time-use data enable the construction of empirically grounded availability profiles that reflect real-world daily routines.

First, individual-level availability is constructed for each 10-minute interval. We define a person as available if they are at home and not engaged in activities that indicate practical inaccessibility to parcel delivery. Specifically, availability is coded if the individual is at home and not sleeping or engaged in outdoor activities such as gardening or walking the dog. This definition captures not only physical presence but also practical availability for personal handover, which is relevant for last-mile delivery operations. Second, availability is aggregated to the household level, as parcel deliveries are typically received by any available household member. For each household and time interval, we define availability as a binary indicator that equals one if at least one household member is available and zero otherwise. This aggregation reflects the operational perspective of delivery processes and results in household-level availability profiles. Third, each household-day is represented

as a 144-dimensional binary vector that captures availability across all 10-minute intervals of the day. To identify typical availability patterns, we apply k-means clustering to these vectors, as this algorithm is simple, fast, and can handle large datasets (Kassambara, 2017).

Prior to clustering, missing values were checked, but none were present. Similar to the approach of Engelhardt (2023), the number of clusters is determined using 30 different methods to identify the best number of clusters via the R Package NbClust, which yields four clusters, as most of the 30 methods indicate four clusters.

The cluster centroids represent archetypical daily availability profiles. To improve usability in vehicle routing algorithms, the 10-minute availability probabilities are aggregated into hourly averages, yielding 24-dimensional profiles for each cluster. Additionally, cluster sizes are calculated as relative frequencies of household-day assignments and used to represent the empirical distribution of availability patterns. The resulting set of four archetypical availability profiles is shown in Figure 1 and captures heterogeneous daily routines, such as full-time working households with low daytime availability (Cluster 4), continuously available households (Cluster 3), and intermediate patterns (Cluster 1 and 2). These empirically derived profiles serve as realistic input for the subsequent simulation experiments.

2.2. Empirical Shipment Data and Instance Generation

To ensure that the experiments reflect realistic last-mile operating conditions, we use empirical shipment data from a German logistics service provider in Bremen. For privacy reasons, the data is not publicly available. The raw data contain operational shipment-level information such as tour identifiers, stop types, delivery addresses, event data, and shipment details. The data are processed and consolidated to obtain a stop-based dataset suitable for generating routing instances.

First, only operational tours relevant to the research objective are retained. Specifically, we keep tours that focus on parcel delivery stops. This ensures that the resulting instances represent actual parcel delivery operations rather than tour types irrelevant to the analysis, such as mail delivery. Next, the shipment-level data are aggregated to the stop level. Because multiple shipment records may belong to the same physical stop, all records associated with the same tour and stop identifier are consolidated into a single stop. During this process, shipment counts are summed, and operational attributes such as stop type and address information are harmonized.

Second, geographic coordinates are assigned to the stops by linking stop addresses to an address reference file containing latitude and longitude information. This results in a geocoded stop dataset that can be used as spatial input for routing optimization. After data cleaning and preparation, the final dataset comprises 1,681 stops from 36 working days, of which 61% are business-to-consumer (B2C) stops. These 36 working days form the empirical basis for instance generation. Third, we generate one base instance per working day, resulting in 36 base instances. Each base instance represents the set of stops that occurred on that day and serves as the routing input for one daily delivery planning problem. The daily structure is intentionally maintained to preserve realistic variation in demand volume, stop composition, and spatial dispersion across operating days.

To analyze the role of incomplete household availability information, we then create multiple scenario variants for each base instance. Customer availability profiles are assigned only to B2C stops, as household availability is relevant for residential deliveries but not for business or depot-related stops. Based on the four empirical availability-profile clusters derived in Section 2.1, each B2C stop is assigned one archetypical profile. This produces the full-information version of each daily instance. To simulate realistic information constraints, the assigned availability profiles are subsequently masked for a defined share of B2C stops. Masking means that the stop remains part of the routing instance, but no customer availability profile is provided to the optimization model for that stop. In this way, the experiments reflect the real-world situation in which the information may be available only for a subset of customers. We consider ten availability scenarios, in which 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, and 100% of B2C stops have a household availability profile. The masking is performed separately for each base instance, yielding daily variants with systematically increasing levels of household availability information.

Overall, this procedure results in a structured experimental dataset comprising 36 empirical base instances and 10 availability-information scenarios per instance, for a total of 396 instances. The resulting instances provide the foundation for analyzing how the degree of coverage of customer availability profiles influences the value of availability-aware last-mile routing. Figure 2 shows the operating area in Bremen and the stop distribution of one instance (operating day). The blue points represent one single stop, and the green and yellow points represent multiple stops in a dense area.

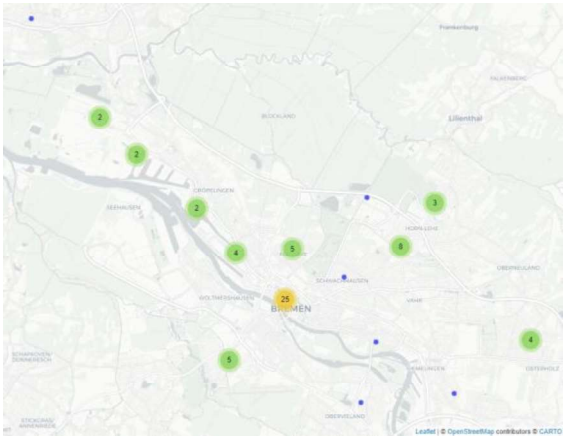

Figure 2: Empirical stop distribution of one instance (one exemplary operating day)

2.3. Vehicle Routing with Household Availability Profiles

To evaluate the operational value of household availability information, we apply the Vehicle Routing Problem with Availability Profiles (VRPAP) proposed by Voigt et al. (2022). In contrast to the classical vehicle routing problem, the VRPAP explicitly incorporates the risk of failed delivery attempts into routing decisions. This is particularly relevant in last-mile delivery settings, where not all customers are equally likely to be available at the time of service.

The key distinction from a conventional VRP lies in the objective function. In addition to routing-related costs, the VRPAP also considers the expected costs of failed deliveries. These expected failure costs depend on two components: the probability that a customer is unavailable at the time of service and the cost associated with an unsuccessful delivery attempt. As a result, the optimization does not solely seek the shortest or cheapest set of tours, but instead balances additional detour and service costs against the potential savings from avoiding failed deliveries.

The optimization model and solution approach are adopted from Voigt et al. (2022), who provide the formal mathematical problem description and the corresponding solution method for this analysis. The problem is solved using a hybrid adaptive large neighborhood search (HALNS) metaheuristic, which is well-suited for this NP-hard problem class and allows the efficient solution of realistically sized empirical instances. Since the vehicle routing model itself is not the contribution of this paper and is only used as an evaluation instrument, these technical details are not repeated here.

The algorithm is applied to the full set of 396 instances generated in Section 2.2. These instances result from combining 36 empirical base instances with 11 availability-information scenarios ranging from 0% (base scenario) to 100% profile availability. For each instance, the solver generates a routing plan that accounts for the household availability profiles, where provided, and falls back to the standard routing logic where such profiles are not available. For each scenario, we compute a set of operational performance indicators based on the following assumptions, validated in discussions with practitioners: service times of 2 minutes per stop, driver costs of 18€ per hour, vehicle costs of 0.34 €/km, and failed delivery costs of 4,50€. The performance indicators include failed delivery rate, delivery-related costs, and total kilometers traveled. Together, these indicators capture the service-related, economic, and environmental implications of availability-aware routing. By comparing results across scenarios, we assess how the degree of household availability information affects last-mile delivery performance.

3. Results and Discussion

3.1. Effect of APs on Failed Deliveries

This section analyzes the impact of household availability profiles (APs) on the occurrence of failed deliveries. Figure 3 illustrates the distribution of daily failed delivery rates across all instances for increasing shares of availability profiles.

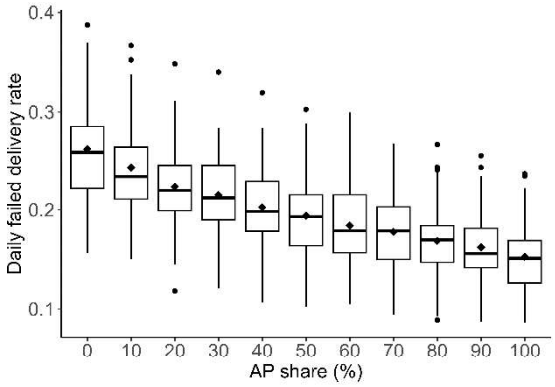

Figure 3: Failed Delivery rates and AP share scenario

The results show a clear, consistent pattern: as the share of customers with availability profiles increases, the failed-delivery rate decreases. This effect is visible across all scenarios and is robust to variation in daily demand structures. Even a relatively small share of availability information leads to a noticeable reduction in failed deliveries, while higher availability coverage further amplifies this effect. More specifically, the median failed delivery rate steadily declines as availability information increases from 0% to 100%.

These findings provide strong empirical evidence for the core value proposition of availability profiles in last-mile logistics. By incorporating customer availability into routing decisions, delivery attempts can be better aligned with the likelihood of successful handover, thereby reducing the number of failed deliveries and improving delivery service quality. The more customers

provide availability data, the greater the potential for improvement.

3.2. Effect of APs on Last-Mile Costs

This section analyzes the impact of APs on last-mile delivery costs. In contrast to the improvements in delivery success rates observed in Section 3.1, the results indicate that tour costs increase as the share of availability profiles rises. Figure 4 shows the distribution of daily tour costs per stop across all AP scenarios.

A clear upward trend can be observed: as the share of customers with availability profiles increases, median costs rise steadily. This effect is consistent across all instances and becomes more pronounced at higher levels of availability coverage. From a conceptual perspective, this result is not surprising.

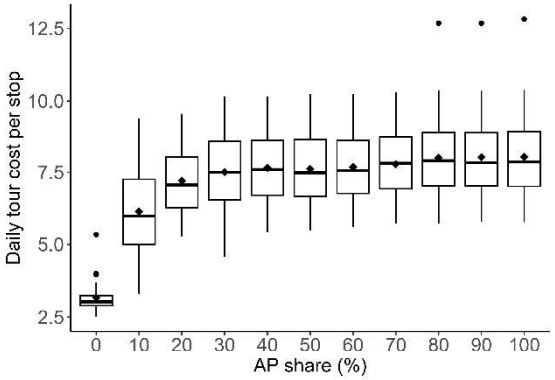

Figure 4: Time- and distance-dependent tour cost per stop and AP share scenario

Incorporating availability profiles into routing decisions requires adjusting routes to better align service times with customer presence, which may introduce additional operational effort. However, a closer examination of the results reveals an important insight. As shown in Chapter 3.3, the total travel distance increases only marginally. This indicates that the observed cost increase is not primarily driven by longer routes, but rather by time-based cost components. The main driver of this effect is methodological. The routing algorithm allows for waiting times at customer locations in order to synchronize delivery attempts with periods of higher availability. Since waiting times are currently not penalized in the cost function, the algorithm may exploit this flexibility excessively. As a result, routes may include substantial idle times, leading to higher total tour durations and thus increased costs, even though the spatial structure of the routes remains largely unchanged.

This finding highlights an important limitation of the current modeling setup rather than a fundamental inefficiency of availability-aware routing. In practical applications, waiting times are typically associated with opportunity costs and therefore incorporated into the objective function. Ongoing work addresses this issue by refining the cost structure so that waiting times are accepted only when economically justified.

Consequently, the currently observed increase in tour costs is expected to decrease significantly under a more realistic cost formulation. In combination with the reduction in failed deliveries demonstrated in Section 3.1 and the associated failed delivery costs, this suggests that the use of availability profiles could become economically beneficial once cost modeling is further improved and total last-mile costs are considered, including distance-based, failed-delivery-based, and all-time-based cost components.

3.3. Effect of APs on Last-Mile Distance

This section analyzes the impact of APs on the total distance traveled in last-mile delivery. In contrast to concerns that availability-aware routing may lead to inefficient routing patterns (‘Zig-Zag-Tours’), the results show that the inclusion of availability profiles does not significantly increase travel distance. Figure 5 presents the distribution of daily distance per stop across all AP scenarios. The boxplots reveal that both the median and dispersion remain largely stable as AP coverage increases. While minor fluctuations can be observed, no systematic upward trend is evident. This indicates that incorporating availability information does not substantially lengthen routes.

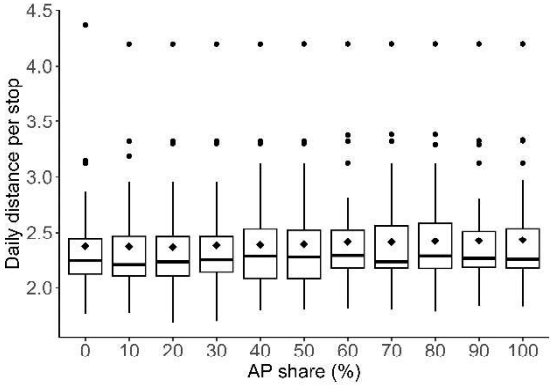

Figure 5: Distance per stop and AP share scenario

To further investigate this relationship, we estimate a simple linear regression model of the share of available APs on distance per stop across all instances. The results suggest that the effect of AP coverage on distance is statistically insignificant ($P > 0{,}05$) and economically negligible.

To complement the quantitative analysis, Figure 6 visually compares routing structures for three representative days under different scenarios (AP00, AP50, and AP100). Availability-profile stops are highlighted in red. These visualizations show that the overall spatial structure of the routes remains largely unchanged. Instead of creating zigzag patterns or inefficient detours, the routing algorithm performs localized resequencing, where individual stops are repositioned within existing

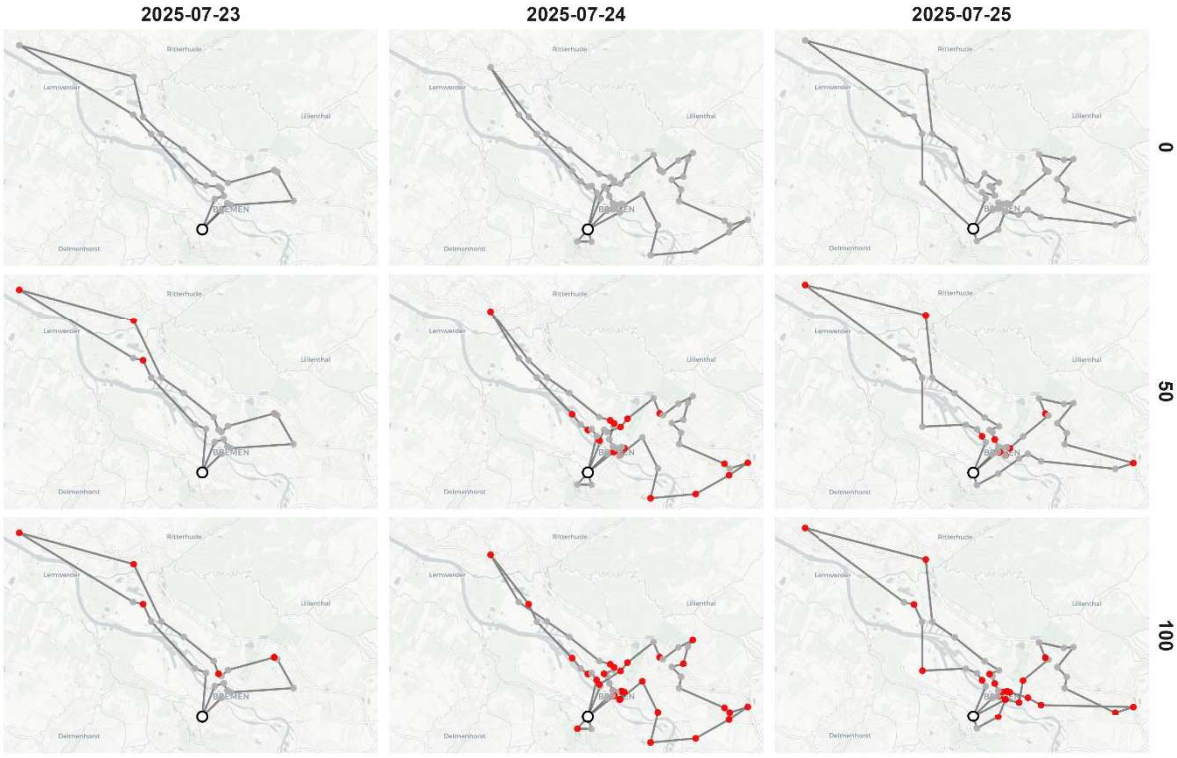

Figure 6: Exemplary operating days with different AP levels; Lines = Tours, Grey

routes in a distance-efficient manner. This confirms that the additional routing complexity introduced by availability constraints remains limited.

The key implication of these findings is that availability-aware routing does not increase travel distance or associated emissions. This challenges the intuitive expectation that incorporating customer availability necessarily leads to less efficient routing patterns. Furthermore, the results suggest that the approach can improve service performance without imposing additional environmental burdens. Moreover, by reducing failed delivery attempts (as shown in Section 3.1), availability-aware routing may even improve system efficiency by avoiding repeated delivery trips and thereby freeing up operational capacity. However, this conclusion should be interpreted with caution. As discussed in Section 3.2, the current modeling setup allows for excessive waiting times, which may influence routing decisions. Ongoing work aims to refine the cost structure, and future research will reassess the effects on distance under improved modeling assumptions.

Overall, the findings indicate that integrating APs enables service improvements without increasing routing distance, highlighting the practical feasibility of the approach in real-world last-mile operations, as in the case under investigation.

4. Limitations and Conclusion

This study investigates the value of household availability profiles (APs) for last-mile routing using empirical data and mathematical optimization-based experiments. The results provide insights into the operational effects of availability-aware routing. The most important finding is that availability profiles significantly reduce failed delivery rates, thereby improving service quality even if not all households share their APs. This confirms the central premise of availability-aware routing: aligning delivery attempts with customer presence increases the likelihood of successful first-time deliveries. At the same time, the results show that routing distances remain largely unchanged, indicating that integrating availability information does not lead to inefficient routing patterns or increased environmental impact. However, total tour costs increase with higher availability coverage, primarily due to time-based effects such as waiting times. This effect is not driven by longer routes but by the current modeling assumption that inadequately penalizes waiting times.

These findings should be interpreted in light of several limitations. First, the empirical shipment data are derived from a single logistics service provider in one urban area (Bremen), which may limit the generalizability of the results to other regions or network structures. Second, the availability profiles are based on time-use survey data, which, although representative, do not capture real customer-specific behavior or dynamic day-to-day variability. Third, the assignment of availability profiles to stops and the masking approach rely on stylized assumptions, particularly regarding the distribution of profile types and the availability of

information, and ignore alternative delivery options such as neighbors, safe places, parcel lockers, and shops. Fourth, the current cost modeling does not fully reflect real-world operational constraints, as waiting times are not explicitly penalized, which likely leads to an overestimation of routing costs under availability-aware planning.

Despite these limitations, the study provides important insights into the applicability of availability profiles in real-world last-mile operations. The results suggest that service improvements can be achieved without increasing travel distance, and that the observed cost increases are likely to decrease once the cost structure is refined. Ongoing work addresses this issue by integrating economically meaningful waiting-time penalties into the optimization model. This is expected to significantly improve the economic performance of availability-aware routing and may even lead to a net-positive business case when combined with reduced failed delivery costs.

Future research should focus on several directions. First, the methodology should be validated using additional datasets from different regions and logistics networks to improve external validity. Second, more granular and dynamic availability data, for example, derived from real customer behavior or digital tracking, could enhance the realism of the profiles. Third, integrating alternative delivery options represents a promising extension. Finally, further refinement of the optimization model, particularly regarding time-dependent costs and operational constraints, is necessary to fully capture real-world decision-making.

From an exploitation perspective, the results indicate strong potential for data-driven decision support systems in last-mile logistics. Availability-aware routing can be integrated into existing planning systems to improve delivery success rates and customer satisfaction. With improved cost modeling, the approach has the potential to become economically viable and scalable, especially in environments where failed deliveries are costly and customer data is increasingly available.

In conclusion, this study demonstrates that customer availability profiles are a promising lever for improving last-mile delivery performance. While further methodological refinements are required, the approach offers a realistic pathway toward more efficient, reliable, and customer-centric last-mile logistics.

References

Boyer, K. K., Prud'homme, A. M. & Chung, W. (2009). *The last mile challenge: Evaluating the effects of customer density and delivery window patterns*. Journal of Business Logistics, 30, 185-201. https://doi.org/10.1002/j.2158-1592.2009.tb00104.x

Bundesverband Paket- und Expresslogistik e. V. (BPEX). (2023). *KEP-Studie 2023 - Analyse des Marktes in Deutschland.* Retrieved on 28.05.2026 from https://bpex-ev.de/files/biek/downloads/papiere/BIEK_KEP-Studie_2023.pdf

Engelhardt, M. (2023). *Who is willing-to-pay for sustainable last mile innovations?* Transportation Research Procedia, 69, 910-917. https://doi.org/10.1016/j.trpro.2023.02.252

Escudero-Santana, A., Cortés-Muñoz, J.-C., Lorenzo-Espejo, A. & Pegado-Bardayo, A. (2026) *The impact of customer habit knowledge on last-mile logistics costs and satisfaction.* Research in Transportation Economics, 117, 101777. https://doi.org/10.1016/j.retrec.2026.101777

Florio, A. M., Feillet, D. & Hartl, R. F. (2018). *The delivery problem: Optimizing hit rates in e-commerce deliveries*. Transportation Research Part B: Methodological, 117, Part A, 455-472. https://doi.org/10.1016/j.trb.2018.09.011

Kassambara, A. (2017). *Unsupervised Machine Learning, Practical Guide To Cluster Analysis in R*. STHD. Retrieved on 28.05.2026 from https://xsliulab.github.io/Workshop/2021/week10/r-cluster-book.pdf

Lim, S. F. W. T., Wang, Q., & Webster, S. (2023). *Do it right the first time: Vehicle routing with home delivery attempt predictors*. Production and Operations Management, 32, 1262-1284. https://doi.org/10.1111/poms.13926

Özarık, S. S., Veelenturf, L. P., Van Woensel, T. & Laporte, G. (2021). *Optimizing e-commerce last-mile vehicle routing and scheduling under uncertain customer presence*. Transportation Research Part E: Logistics and Transportation Review, 148, 102263. https://doi.org/10.1016/j.tre.2021.102263

Özarık, S. S., Lurkin, V., Veelenturf, L. P., Van Woensel, T. & Laporte, G. (2023). An Adaptive Large Neighborhood Search heuristic for last-mile deliveries under stochastic customer availability and multiple visits. Transportation Research Part B: Methodological, 170, 194-220. https://doi.org/10.1016/j.trb.2023.02.016

Pan, S., Giannikas, V., Han, Y., Grover-Silva, E. & Qiao, B. (2017). *Using customer-related data to enhance e-grocery home delivery*. Industrial Management & Data Systems, 117(9), 1917-1933. https://doi.org/10.1108/IMDS-10-2016-0432

Peppel, M., Ringbeck, J., & Spinler, S. (2022). *How will last-mile delivery be shaped in 2040? A Delphi-based scenario study*. Technological Forecasting and Social Change, 177, 121493. https://doi.org/10.1016/j.techfore.2022.121493

Praet, S. & Martens, D. (2020). *Efficient Parcel Delivery by Predicting Customers' Locations*.

Decision Sciences, 51, 1202-1231. https://doi.org/10.1111/deci.12376

Schröder, J., Heid, B., Neuhaus, F., Kässer, M., Klink, C. & Tatomir, S. (2018). *Fast forwarding last-mile delivery – implications for the ecosystem.* Retrieved on 28.05.2026 from https://www.mckinsey.com/~/media/mckinsey/industries/travel%20logistics%20and%20infrastructure/our%20insights/technology%20delivered%20implications%20for%20cost%20customers%20and%20competition%20in%20the%20last%20mile%20ecosystem/fast-forwarding-last-mile-delivery-implications-for-the-ecosystem.pdf

Seghezzi, A., Mangiaracina, R. (2023). *Smart home devices and B2C e-commerce: a way to reduce failed deliveries*. Industrial Management & Data Systems, 123(5), 1624-1645. https://doi.org/10.1108/IMDS-10-2022-0651

Statistisches Bundesamt. (2024). *Zeitverwendungserhebung 2022.* Retrieved on 28.05.2026 from https://campus-file-fdz.nrw.de/index.php?strAction=user.datasets

van Duin, J. H. R., de Goffau, W., Wiegmans, B., Tavasszy, L. A. & Saes, M. (2016). *Improving Home Delivery Efficiency by Using Principles of Address Intelligence for B2C Deliveries*. Transportation Research Procedia, 12, 14-25. https://doi.org/10.1016/j.trpro.2016.02.006

Voigt, S., Frank, M., Fontaine, P. & Kuhn, H. (2022). *The Vehicle Routing Problem with Availability Profiles*. Transportation Science, 57(2), 531-551. https://doi.org/10.1287/trsc.2022.1182