Selecting the most promising areas to introduce pooled feeder services when the demand is unknown: case study of Krakow, Poland

Olha Shulika*[1], Hanna Vasiutina[1], Michal Bujak[2, 3], Farnoud Ghasemi[2,3], and Rafal Kucharski[4]

[1]Assistant Professor, Faculty of Mathematics and Computer Science, Jagiellonian University, Poland
[2]Assistant, Faculty of Mathematics and Computer Science, Jagiellonian University, Poland
[3]PhD Candidate, Doctoral School of Exact and Natural Sciences, Jagiellonian University, Poland
[4]Professor, Faculty of Mathematics and Computer Science, Jagiellonian University, Poland

Short summary

This study proposes a parameter-free method to assess the potential of pooled on-demand transit feeder services in urban areas with unknown demand. We introduce the 'fraction of demand', reflecting the probability that a resident will use the service. Demand is generated on the distribution of residents' address points at varying demand fraction levels. Through simulations, we match travellers into pooled rides and evaluate the service's potential using three performance indicators (KPIs). We observe how these KPIs change with varying demand fractions and identify the most promising hub for each area. By setting KPI thresholds, we select the optimal combination of area and hub that meets these thresholds at the lowest demand fraction. This approach assists municipalities in selecting the best areas to launch new services despite the absence of exact demand data. We illustrate its application through a case study in Krakow, ranking 12 pre-selected areas for feeder service deployment.
Keywords: shared mobility, ride-pooling, on-demand feeder, ride-sharing.

1 Introduction

Cities worldwide aim to reduce car dependency and enhance urban sustainability by promoting public transport (PT) and shared mobility services. Addressing first-mile challenges, connecting residential areas with transit hubs, is crucial to encouraging PT usage and decreasing reliance on private vehicles. Flexible on-demand systems, such as demand-responsive transit (DRT), offer adaptable and effective solutions, particularly in low-demand areas where traditional fixed-route services often fail. Ride-pooling, a specific subset of DRT, optimises operations by pooling passengers travelling in similar directions, reducing costs, travel times, and environmental impact (Alonso-Mora et al., 2017). However, these systems face challenges such as the lack of demand data CIVITAS (2024), public resistance (Alonso-González et al., 2020), and funding constraints (Vuchic, 2017). Additionally, regulatory barriers and the need for equitable service distribution complicate their implementation (Shaheen & Cohen, 2020).
Krakow, Poland, is attempting to address its first-mile challenges by implementing a flexible on-demand feeder system. The city plans to introduce on-demand buses to connect residents in one of 12 candidate low-density areas (Fig. 1) to high-frequency tram and train hubs. The proposal involves dispatching on-demand small-capacity buses during the morning rush hour with designated pick-up points feeding into transfer hubs (Fig. 2). From the hubs, travellers continue their journeys using the city's efficient public transport network. This feeder service aims to improve accessibility and reduce private vehicle usage. However, with uncertain demand, the city needs a method to determine which areas to prioritise for service launch. Designing on-demand feeder systems is challenging due to the lack of reliable demand data. Traditional methods rely on parameter tuning, which often misaligns with actual ridership. To overcome this, we propose a parameter-free approach that uses probabilistic inputs, such as population density, to simulate demand fractions and evaluate KPIs such as vehicle mileage reduction, passenger comfort, and occupancy. This method identifies areas with the greatest potential for efficiency at minimal demand levels, ensuring better alignment between the planned and actual performance of the service.

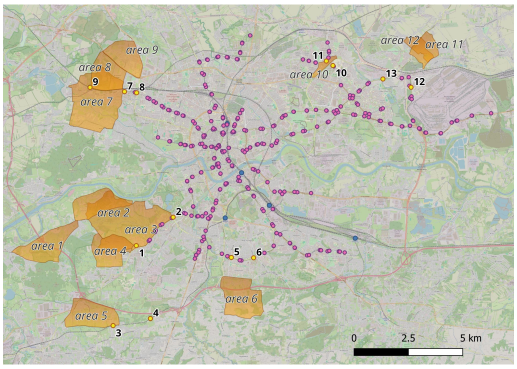


Figure 1: Krakow pre-selected areas (in orange) with corresponding hubs (in yellow), tram stops (in pink) and train stops (in blue).

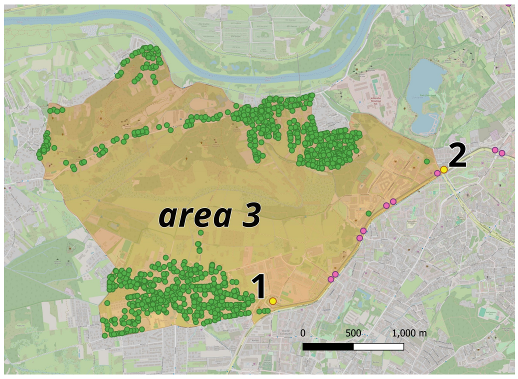


Figure 2: Example of the spatial distribution of address points (in green), tram stops (in pink) and light rail hubs (in yellow) for the Area 3.

To assess which area has the greatest potential before the service is introduced, we identify those that:

- require a lower level of demand to achieve efficiency;
- reach stability at the lowest demand level in performance indicators, marking the required threshold;
- perform better at expected demand levels (if known).

Areas requiring the lowest resident interest to meet efficiency thresholds are prioritised. After selecting the most promising combination, benchmarks are established for service initiation.
By applying this method in a case study of Krakow, we analyse 12 pre-selected areas and rank them based on their potential of implementing on-demand feeder services. This approach helps municipalities in the absence of data on exact demand for a new service compare and select the most promising area to launch the service.

2 Methodology

Our approach to selecting the preferred area to implement on-demand pooled transit feeders integrated with public transport (PT) is illustrated in Fig. 3. Our methodology involves data collection (population distribution, road networks, and hub locations) for pre-selected areas. We consider on-demand bus service as ride-pooling service, a specific subset of DRT, matches passengers travelling in similar directions, optimising operational costs, travel times, and environmental benefits (Alonso-Mora et al., 2017). Therefore, we apply the Exact Matching of Attractive Shared Rides (ExMAS) algorithm (Kucharski & Cats, 2020) to match travellers to the pooled rides and evaluate ride-pooling potential using three key indicators (based on Shulika et al. (2024b)): mileage reduction, passenger satisfaction, and occupancy.
Initially, we assess the progression of these KPIs across varying demand fractions, identifying the most promising hub within each pre-selected area. Following this, we compare the candidate areas by identifying the proportion of residents (fraction of demand, α) who must be interested in the service, denoted as the level α, to meet the following minimum efficiency thresholds required to launch the new service:

- $\triangle T_v$(vehicle hours reduction) ≥ 0.1: the launching of shared rides (instead of individual ones) allows for a reduction of vehicle kilometres by at least 10%;
- $\triangle U_p$(travellers utility gains) ≥ 0.025: passenger comfort improves by at least 2.5% compared to individual travel. For the analysed scenario of free on-demand bus service, this measure ensures that passengers do not encounter significant discomfort associated with a new service;
- O(occupancy) ≥ 2: the average vehicle occupancy exceeds 2.

Areas that meet these thresholds at the lowest demand levels are considered most favourable. Benchmarks provide insights into minimum demand, demand growth points, and consistent KPI

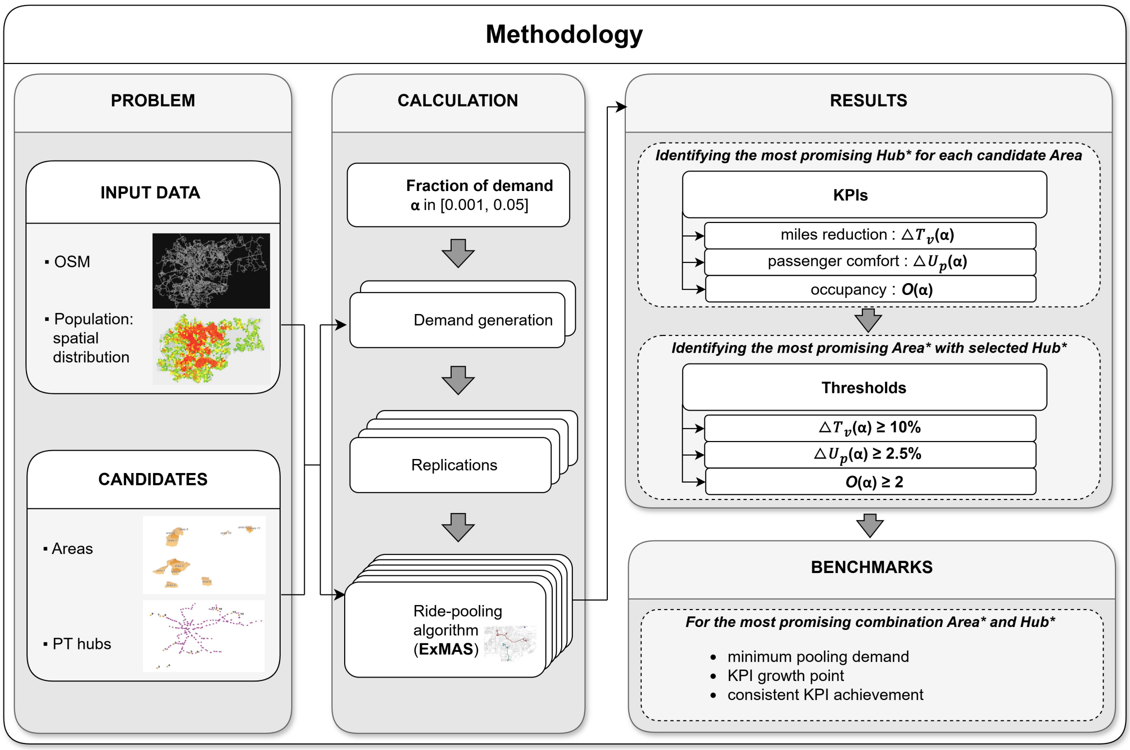


Figure 3: An overview of the applied methodology for selecting a preferred area to implement on-demand pooled transit feeders.

performance levels. The case study demonstrates the application of our method in Krakow, evaluating 12 areas for the implementation of services. Our findings highlight the importance of identifying zones where ride-pooling can thrive at minimal initial demand. Scripts for reproducible results are available in the public repository (Shulika et al., 2024a).

Input

The method relies on OpenStreetMap (OSM) data for the city's road network, population distribution (coordinates of residential address points), and candidate areas near public transport hubs provided by analysts. The non-deterministic approach uses random selection of travellers to analyse results across multiple replications, ensuring meaningful insights through aggregated KPIs.

Demand generation

For each candidate area A, the total population is used to estimate the maximum potential demand. Demand fractions range from 0.1% to 5%, representing minimum feasible pooling to stable KPI levels. Travel requests $\{q_i = (o_i, d, \tau_i)\}$ are generated, where o_i represents origins, d is the nearest public transport hub, and τ_i denotes travel time sampled uniformly within the simulation period. If multiple hubs exist, the demand is split accordingly (Fig. 4). The process repeats for each demand fraction and replication.

ExMAS

ExMAS, an open-source Python algorithm, optimally matches travellers to pooled rides while minimizing mileage (Kucharski & Cats, 2020). To assess whether a pooled ride candidate r_k is attractive to the traveller i, we compare shared rides to private rides using utility formulas:

$$\begin{aligned} U_i^{ns} &= \beta_c \lambda l_i + \beta_t t_i \\ U_{i,r_k}^{s} &= \beta_c (1-\lambda_s)\lambda l_i + \beta_t \beta_s (\hat{t}_i + \beta_d \hat{t}_i^p) + \varepsilon, \end{aligned} \tag{1}$$

where U_{i,r_k}^s, U_i^{ns} denote, respectively, the utility of shared ride (for i-the traveller, ride r_k) and the utility of non-shared ride (for i-th traveller). λ stands for the per-kilometre fare, while λ_s denotes

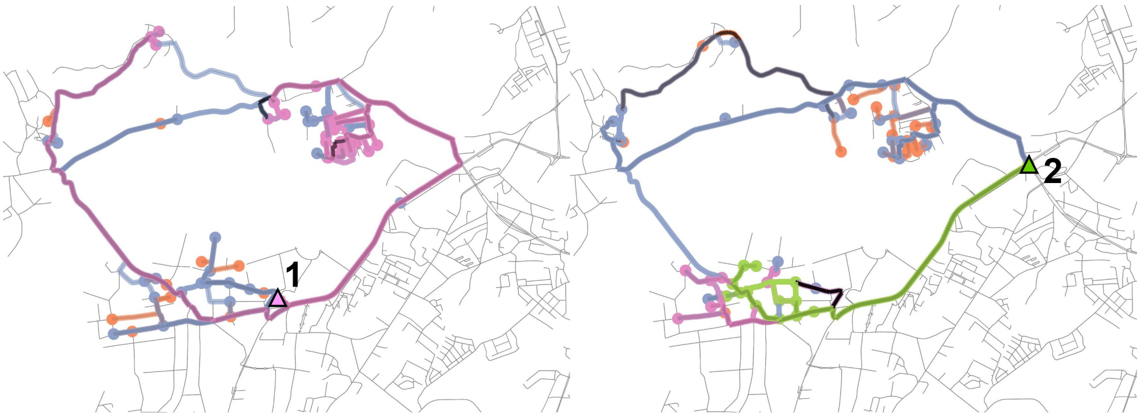


Figure 4: An example of visualising ride-pooling algorithm shows sample rides for Area 3: all sampled travellers of an area are heading from origins (dots) to hub 1 (left pink triangle denoted 1) or to hub 2 (right green triangle denoted 2) as their transit destination points (hubs).

a discount for sharing a ride. β^c, β^t, β^s, β^d are the exogenous behavioural parameters: cost sensitivity, value of time, willingness-to-share and delay sensitivity, respectively. t_i and $\hat{t}_i$ stand for the travel time of non-shared and shared rides, respectively, $\hat{t}_i^p$ is a pick-up delay associated with pooling, and ε is a random term (for the sake of simplicity, assumed here to be constant and null, which yields a deterministic formulation). It assumes that pooling is chosen only when it offers better utility than solo travel, considering costs (fares, delays, detours) and benefits (reduced costs). For each traveller, utilities of shared and nonshared rides are compared, ensuring that each traveller is assigned uniquely. ExMAS explores all ride combinations based on discrete choice theory, but does not explicitly model vehicle fleets.

Performance indicators

We report three KPIs of interest (for full definitions of possible KPIs we refer to Shulika et al. (2024b)). For each combination of area A, hub H and a single realisation of demand with given fraction of demand α, we report:

- what is the potential mileage reduction $\Delta T_v(\alpha, A, H)$;
- how is the (perceived) traveller utility improved $\Delta U_p(\alpha, A, H)$;
- what is the average occupancy $O(\alpha, A, H)$.

Potential mileage reduction and passenger comfort are defined by comparing vehicle hours and traveller utility when the ride-pooling service is available and when it is not applied. Occupancy represents the ratio of total passenger hours in the solo ride-hailing scenario to total vehicle hours in the pooled scenario (Shulika et al., 2024b).

Thresholds

For each area A and hub H, thresholds for KPIs are:

$$\alpha^*_{\Delta T_v}(A, H) = \min_{\alpha \in [0.001, 0.05]} \Delta T_v(\alpha, A, H) \geq 0.1, \tag{2a}$$

$$\alpha^*_{\Delta U_p}(A, H) = \min_{\alpha \in [0.001, 0.05]} \Delta U_p(\alpha, A, H) \geq 0.025, \tag{2b}$$

$$\alpha^*_{O}(A, H) = \min_{\alpha \in [0.001, 0.05]} O(\alpha, A, H) \geq 2.0. \tag{2c}$$

The goal is to identify areas and hubs surpassing these thresholds at the lowest α.

Hub and Area Selection

For each area, the hub achieving the best KPI performance becomes its optimal hub (H^*). The most promising area (A^*) is determined by identifying combinations that meet KPI thresholds with the lowest demand α.

Benchmarks

After selecting the most promising combination (A^*, H^*) we run ExMAS simulations across demand fractions $\alpha \in [0, 0.01]$ in 30 replications. This allows us to explore the most unstable phase, which typically occurs during the initial launch when demand is at its lowest. We establish the following benchmarks:

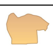
- minimum pooling demand: the minimum fraction of demand α required to successfully pool travellers into shared trips;
- KPI growth point: the demand fraction α at which ride-pooling potential starts to grow;
- consistent KPI achievement: the demand fraction α at which a stable level of all three KPIs is reached.

These benchmarks help set realistic expectations for performance in a specific early stage of service.

3 Results and discussion

The method was tested in 12 areas of Krakow, Poland, paired with nearby transport light rail hubs. Inputs include OSM road network data, municipal population distribution, and hub locations. Areas vary in population density and hub proximity (Table 1). For each area-hub pair demand fractions α range from 0.001 to 0.05, with detailed benchmarks assessed for α values from 0 to 0.01. Simulations replicate demand generation 10 times for area-hub pairs and 30 times for the most promising combinations. The parameters include a six-passenger vehicle capacity, a 30-minute morning peak simulation window, and cost/travel behaviour metrics based on local data.

Table 1: Characteristics of pre-selected areas.

Name	Surface	Population [residents]	Density $[\frac{residents}{km^2}]$	Hub Name	Hub Type	Hub Avg. distance to hub $[km]$
Area 1		1263	489.8	1.'Czerwone Maki P+R'		6.183
Area 2		1286	585.6	1.'Czerwone Maki P+R' 2.'Norymberska'		3.385 5.416
Area 3		4550	676.9	1.'Czerwone Maki P+R' 2.'Norymberska'		1.762 3.454
Area 4		3719	1781.2	1.'Czerwone Maki P+R'		1.744
Area 5		1593	658.6	1.'Czerwone Maki P+R' 3.'Krakow Sidzina' 4.'Krakow Opatkowice'		4.369 1.634 4.163
Area 6		2396	851	5.'Kurdwanów P+R' 6.'Nowosadecka'		1.953 2.077
Area 7		5651	676.9	7.'Bronowice Małe' 9.'Kraków Mydlniki(PKP)'		2.309 1.553
Area 8		1836	904.5	7.'Bronowice Małe' 9.'Kraków Mydlniki(PKP)'		1.072 1.924
Area 9		4002	2202.2	7.'Bronowice Małe' 8.'Bronowice SKA' 9.'Kraków Mydlniki(PKP)'		1.662 1.863 2.911
Area 10		3069	9880.8	10.'Dunikowskiego' 11.'Rondo Piastowskie'		0.503 0.374
Area 11		1925	1586.4	13.'Wańkowicza' 12.'Zajezdnia Nowa Huta'		3.243 1.916
Area 12		941	1334.3	13.'Wańkowicza' 12.'Zajezdnia Nowa Huta'		3.084 1.879

The method effectively selects optimal hubs and areas. Fig.5 exemplifies the stage of the proposed methodology for selecting the preferred hub within each area, using Area 3 as an example. Similar analyses are conducted for all candidate areas, enabling tailored recommendations for each. The preferred hubs for all areas are summarised in Table 2.

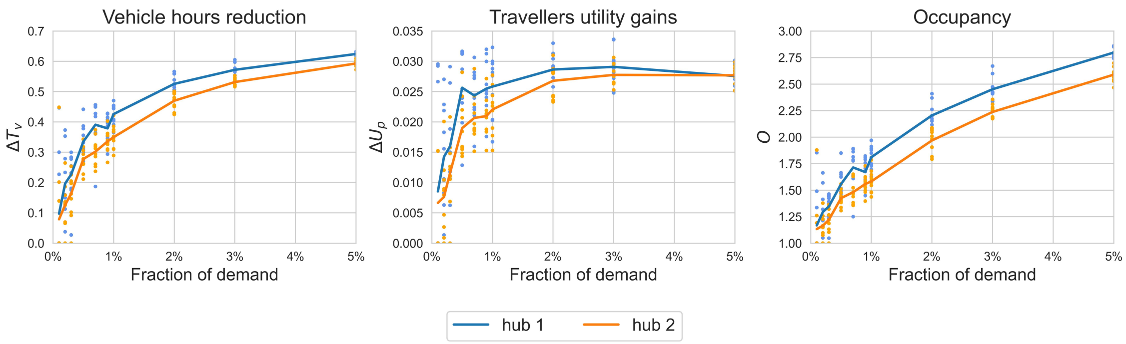


Figure 5: Three key performance indicators of ride-pooling plotted against the fraction of demand for Area 3. The lines represent the average performance across multiple simulations, while the dots represent individual simulation results. Both hubs in Area 3 show similar trends, but hub 1 has a slight edge in potential.

Table 2: Ranking of candidate areas.

Area	**Hub**	**Threshold**						**Total KPI rank score**	**Final area rank**
		$\Delta T_v \geq 0.1$		$\Delta U_p \geq 0.025$		$O \geq 2$			
		α	**rank**	α	**rank**	α	**rank**		
1	1.'Czerwone Maki P+R'	0.005	11	0.01	4	0.02	4	29	10
2	2.'Norymberska'	0.002	4	0.02	7	0.02	4	15	6
3	1.'Czerwone Maki P+R'	0.001	1	0.009	2	0.01	2	5	2
4	1.'Czerwone Maki P+R'	0.003	10	0.05	11	-	11	32	11
5	1.'Czerwone Maki P+R'	0.002	4	0.009	2	0.01	2	8	3
6	6.'Nowosadecka'	0.002	4	0.01	4	0.02	4	12	5
7	9.'Kraków Mydlniki (PKP)'	0.001	1	0.01	4	0.02	4	9	4
8	9.'Kraków Mydlniki (PKP)'	0.002	4	0.02	7	0.02	4	15	6
9	9.'Kraków Mydlniki (PKP)'	0.001	1	0.005	1	0.007	1	3	1
10	10.'Dunikowskiego'	0.02	12	-	12	-	11	35	12
11	12.'Zajezdnia Nowa Huta'	0.002	4	0.02	7	0.02	4	15	6
12	12.'Zajezdnia Nowa Huta'	0.002	4	0.02	7	0.03	10	21	9

Fig.6 shows the ride-pooling potential in pre-selected areas and their preferred hubs. Not all areas reach the thresholds. In particular, Area 10 fails to meet the second and third thresholds, and Area 4 only misses the third threshold. For Area 9, even at a demand level of up to 3%, the thresholds are exceeded. This observation is confirmed by ranking data (Table 2). Ranking the areas according to the minimum demand fraction required to meet three performance thresholds, Areas 3, 7, and 9 achieve the lowest fractions and rank highest, while Area 10 requires a higher fraction and ranks lowest. The total ranking indicates Area 9 paired with Hub 9 'Krakow Mydlniki (PKP)' as the most promising, with a score of 3. Area 3 with Hub 1 'Czerwone Maki P+R' ranks second, while Area 10 with Hub 10 'Dunikowskiego' ranks last.
For Area 9, we evaluate three benchmarks at demand fractions $\alpha \in [0, 0.01]$ over 30 replications (Fig. 7). The results reveal that the pooled rides emerge at a demand fraction of 0.025%. A significant growth in ride-pooling efficiency occurs at 0.05%, marking the point where the service becomes more efficient. The third benchmark, where KPIs consistently meet thresholds, is reached at demand fractions of 0.1%, 0.5%, and 0.7% for three KPIs, respectively. These benchmarks provide the municipality with valuable projections on service efficiency and sustainability during the early phases of implementation.

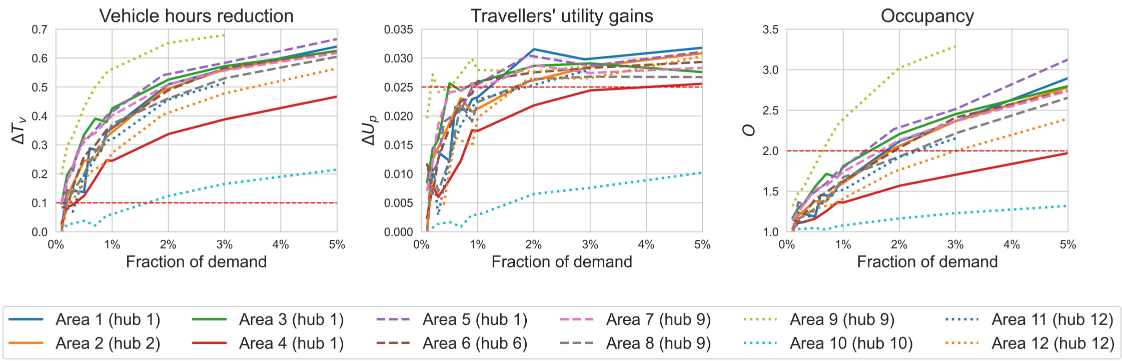


Figure 6: Three key performance indicators of ride-pooling plotted against the fraction of demand for pre-selected areas and the most promising corresponding hubs. Horizontal dashed red lines represent the set thresholds.

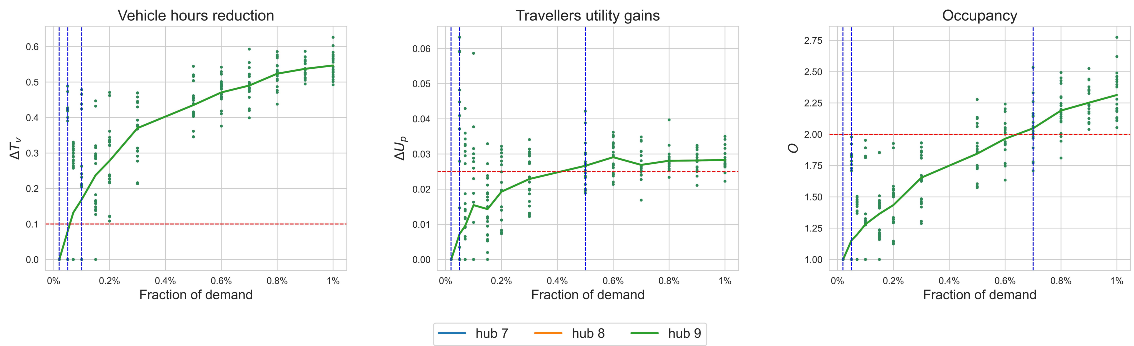


Figure 7: KPIs and three benchmarks for the combination of Area 9 and Hub 9 'Kraków Mydlniki (PKP)', plotted against demand levels. Each dot represents an individual simulation results, while the lines show average performance. Horizontal dashed red lines indicate established KPI thresholds, while vertical blue lines mark three benchmarks.

4 Conclusions

We propose a parameter-free approach using demand fractions to simulate potential demand and evaluate on-demand transit feeders in urban multimodal networks. This method helps municipal authorities compare areas for service launch, even without exact demand data. The Krakow case study demonstrates the method's effectiveness in identifying areas with high potential for on-demand feeder services. Using three KPIs: vehicle kilometre reduction, passenger comfort, and vehicle occupancy, we evaluate the feasibility of service implementation in 12 areas. Area 9, paired with Kraków Mydlniki (PKP), ranks highest (Table 2), tending to have a favourable balance of population density, distance to hubs, and infrastructure suitable for pooled transit (Table 1). In contrast, lower-ranking areas, such as Area 10, with a high population density but close proximity to the Dunikowskiego hub, demonstrate the least potential. We establish benchmarks for service performance in Area 9 with Hub 9. These benchmarks reveal that ride-pooling efficiency improves significantly at 0.05% demand, with efficiency thresholds met at demand fractions 0.1%, 0.5%, and 0.7% for three KPIs, respectively. This information helps municipal planning by forecasting the growth and sustainability of the service. The study findings suggest the viability of using a parameter-free approach to assess the potential for on-demand transit feeders in urban multimodal networks., providing municipalities with insights on optimal locations for service implementation.

Limitations and Future Work

Despite its advantages, this study has limitations. The ExMAS algorithm only evaluates point-to-point ride-hailing, comparing it to solo ride-hailing. Furthermore, demand must be predetermined, and the fleet is not explicitly handled. The study is based on Krakow, a mid-sized European city, and focusses on the first mile segment from pick-up points to hubs, without considering the entire trip. Additionally, we assume that everyone who wants to travel by a given transport system wants to reach one hub. We only consider population potential, without taking into account other factors such as motivation for travel or current transportation habits. Future research should consider the entire traveller trip, including the public transit segment from the hub to the final destination. It should also take into account demographic factors and time-of-day variations to improve demand estimation. Further studies could test the method in various urban contexts to identify universal patterns and better understand how on-demand feeder services can complement traditional public transit.

Acknowledgements

This research was co-funded by the European Union's Horizon Europe Innovation Action under grant agreement No. 101103646.

References

Alonso-González, M. J., Hoogendoorn-Lanser, S., van Oort, N., Cats, O., & Hoogendoorn, S. (2020). Drivers and barriers in adopting mobility as a service (maas)–a latent class cluster analysis of attitudes. *Transportation Research Part A: Policy and Practice*, *132*, 378–401.

Alonso-Mora, J., Samaranayake, S., Wallar, A., Frazzoli, E., & Rus, D. (2017). On-demand high-capacity ride-sharing via dynamic trip-vehicle assignment. *Proceedings of the National Academy of Sciences*, *114*(3), 462–467.

CIVITAS. (2024). *Civitas website.* `https://https://civitas.eu/`. ([Online])

Kucharski, R., & Cats, O. (2020). Exact matching of attractive shared rides (exmas) for system-wide strategic evaluations. *Transportation Research Part B: Methodological*, *139*, 285–310.

Shaheen, S., & Cohen, A. (2020). Mobility on demand (mod) and mobility as a service (maas): Early understanding of shared mobility impacts and public transit partnerships. In *Demand for emerging transportation systems* (pp. 37–59). Elsevier.

Shulika, O., Bujak, M., Ghasemi, F., & Kucharski, R. (2024a). *Exmas-plotting.* `https://github.com/OlhaShulikaUJ/SUM_project/tree/main/NSM`. ([Online])

Shulika, O., Bujak, M., Ghasemi, F., & Kucharski, R. (2024b). Spatiotemporal variability of ride-pooling potential–half a year new york city experiment. *Journal of Transport Geography*, *114*, 103767.

Vuchic, V. R. (2017). *Urban transit: operations, planning, and economics.* John Wiley & Sons.