Scaling Urban Logistics Microsimulation: Traffic Analysis Zone Similarity Assessment to Estimate Simulation Results

Emilian Szczepański[1*], Jan Paszkowski[2], Marianna Jacyna[3], Kai Wang[4], and Jiajing Gao[5]

[1, 2, 3] Warsaw University of Technology, Warsaw 00-662, Poland
[4] Tsinghua University, Beijing, 100084, China
[5] Shanghai University, Shanghai, 200444, China
*emilian.szczepanski@pw.edu.pl

Abstract. This article presents a method of assessing a similarity of the traffic analysis zones to the given base zone. For the research purposes, three variants of data to compare similarity have been used. First variant of are the GIS urban development data combined with the travel motivation data. Second are the energy consumption data for the available vehicle types and third variant are the post infrastructure data, such as post offices and parcel lockers. For the validation of the similarity, the chosen twin-similar zones have been tested in the microsimulation model in Flexsim software. In this model, a package delivery process have been simulated in three scenarios. First is a package delivery for the specific address, second is the delivery to the package locker and the third is the delivery by the parcel locker directly to the address by the delivery drones. The results of the simulation will be compared between variant zones and the reference zone to discover which variant of the similarity data are most suitable for the delivery traffic reproduction.

Keywords: Freight Traffic Modelling, Traffic Analysis Zone Similarity, Urban Logistics.

1 Introduction

The development of cities and changes in consumer behavior significantly influence the shaping of urban logistics systems. Modern technologies and transport solutions, the growing share of e-commerce, and the expansion of parcel locker networks contribute to increasing delivery efficiency while considering customer needs. However, various barriers often arise, such as traffic congestion, low-emission zones, or other restrictions resulting from local policies, including weight or time-based access limitations to specific urban areas. The efficiency of urban freight delivery systems depends on multiple factors. Its improvement can be achieved through various distribution strategies, such as integrating deliveries with public transport vehicles [1], utilizing drones or robots for distribution purposes [4], [11], [12] or employing intermodal infrastructure

X. Qu et al. (eds.), *Proceedings of the 14th Asia-Pacific Conference on Transportation and the Environment (APTE 2025)*, Atlantis Highlights in Engineering 41,
https://doi.org/10.2991/978-94-6463-972-8_4

in last-mile logistics [7], [8], [14]. Additionally, different configurations and specific strategies can be applied to various transport modes and transshipment infrastructures. The influence of modern solutions referred to as CAV (Connected Automated Vehicles) should also be considered, as well as the impact of communication and dynamic control of these vehicles on distribution processes [13], [17]. The implementation of such solutions presents a significant challenge due to differences in urban areas, including building density, population size, wealth levels, as well as the characteristics of recipients, suppliers, and the road network. Additionally, municipal transport policies further influence implementation feasibility. Each city—and even different areas within a single city—has unique characteristics, making experimental deployments without prior simulation studies less likely to succeed. Failure in such cases may hinder the development of environmentally friendly, consumer and business-supportive solutions.

Simulation studies at the macro scale for larger areas within a systemic approach are a standard practice in planning public transport systems and transport infrastructure. They are used to analyze the distribution of traffic flow across the road network and assess related costs and emissions. At the micro scale, simulation studies are employed to examine specific components of the transport system and evaluate particular solutions with a higher level of detail. This often involves agent-based modeling or the representation of decision-making problems in dynamic environments. [5].

It is therefore necessary to emphasize that the implementation of specific solutions in a given city requires linking both micro- and macro-scale studies. However, such studies are costly due to the labor-intensive nature of the research and the difficulty of obtaining detailed data necessary for accurately representing both the urban system and the freight transport system.

This article presents a concept based on the theory of similarity between zones to reduce the workload associated with simulation studies of implementing modern and sustainable solutions in urban logistics. An important issue in this context is the delineation and characterization of zones within a city [2], [10], [16], Properly defined and described zones with a uniform structure enable rational traffic planning and analysis. At the same time, they create the potential for comparing different solutions and transferring them to other areas.

The issue of similarity analysis of traffic analysis zones has been explored from various perspectives. McKenzie and Romm [9] used cosine similarity in micromobility traffic to identify similar areas in two cities, aiming to compare traffic patterns regardless of traffic volume. Dong et al. [3] applied Euclidean vector similarity analysis to distinguish residential and commercial zones in urban traffic zone division. Kim [6] employed a novel dynamic time warping distance to measure urban similarities using smart card data. Based on similarity theory, Wang et al. [15] proposed a method for supplementing missing data in models.

The aim of this article is to verify the possibility of identifying zones with similar characteristics within a given city and approximating the micro-simulation results from a base zone to these similar zones. To achieve this, a macro-scale model developed in PTV VISUM was used as a data source, while a micro-scale model implemented in Flexsim served as the simulation environment for the delivery system. The outcome of the study is an analysis of the quality of this approximation and guidelines for simplifying large-scale simulation studies.

2 Research method

This study aims to verify the preliminary assumptions regarding the feasibility and relevance of such an approximation. Future research will focus on analyzing the impact of various factors on these approximation models.

To achieve this goal, a research methodology was developed, as presented in Figure 1. To maintain consistency in the approach and establish a connection between micro- and macro-simulation models, the macroscopic simulation model developed in PTV VISUM and the GIS database (BDOT10k for spatial planning) were used as the primary data source and starting point. The model provides data on the division of the area into traffic analysis zones as well as the characteristics of each zone, including demographic data, travel motivations, transport infrastructure, and building data from the GIS database. Additionally, it allows for traffic distribution analysis across the urban network, enabling the determination of transport system energy demand and traffic intensity for each zone.

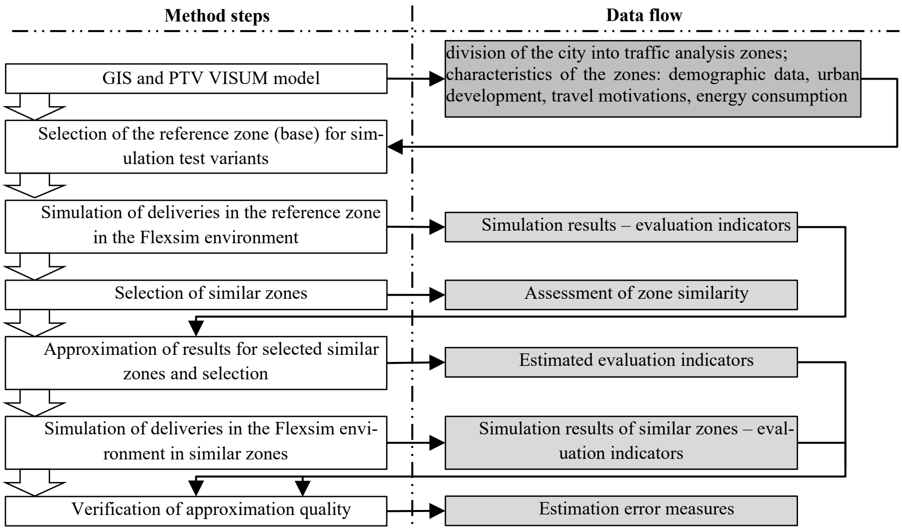

Fig. 1. Diagram of the research approach

3 Case study – areas in GZM metropolis, Katowice, Poland

3.1 Assumptions, variants and scenarios description

The study area for which the analysis of the feasibility of approximating simulation studies across zones was conducted is located in Poland, within the Silesian region.

GZM Metropolis is an urban agglomeration consisting of 41 cities and municipalities. The population of this area is approximately 2.3 million inhabitants, and its total area covers 2,500 km². The largest city and central hub of the region is Katowice. For this area, a transport model in PTV Visum environment was obtained. Based on this model, an analysis of the courier parcel delivery system was conducted, with a particular focus on the city of Katowice.

For the conducted analyses, one base and three alternative research variants were established, each considering different characteristics in assessing zone similarity: *V0* – a base variant determined through expert evaluation, *V1* – based on spatial data and travel motivations, *V2* – based on energy consumption, *V3* – based on the density of postal infrastructure. Similarity analysis was performed among the four selected zones to compare data vectors between the reference zone and the tested variants. In this paper among the data vector comparison methods, an Euclidean distance metric have been used - formula (1), for a straightforward results analysis. In the future works, other methods will be used for the more precise analyses (i.e. Mahalanobis).

$$\forall i, j \quad d(\boldsymbol{A}_{Vi,Z1}, \boldsymbol{A}_{ViZj}) = \sqrt{\sum_{k=1}^{p(Vi)} \left(A_{Vi,Z1}(k) - A_{Vi,Zj}(k) \right)^2} \tag{1}$$

where: i – the index of the variant and $i \in \{1, 2, 3\}$, it was assumed that *V0* is a base variant for which the reference zone is designated as *Z1*; j – the index of the zone, $j \in \{1, 2, \ldots, 88\}$; $\boldsymbol{A}_{Vi,Z1}$, $\boldsymbol{A}_{Vi,Zj}$ – characteristic vectors of the reference zone *Z1* and the compared zones within a given variant; $A_{Vi,Z1}(k)$, $A_{Vi,Zj}(k)$– the characteristics of a given zone, where k corresponds to the characteristic index, are presented respectively for the reference zone and the compared zones. For each variant, the characteristics (k) and their number ($p(Vi)$ differ depending on the specific approach applied in the similarity analysis:

- for *V1* the characteristics $A_{V1,Zj}(k)$ include:
  - $k = 1$: Office buildings area (for all floors) [sqm],
  - $k = 2$: Service and shopping buildings area (for all floors) [sqm],
  - $k = 3$: Low density housing area (for all floors) [sqm],
  - $k = 4$: High density housing area (for all floors) [sqm],
  - $k = 5$: University area (for all floors) [sqm],
  - $k = 6$: Industrial buildings area (for all floors) [sqm],
  - $k = 7$: Hospital area (for all floors) [sqm],
  - $k = 8$: Transportation-related buildings area (for all floors) [sqm],
  - $k = 9$: Warehouse and storage buildings area (for all floors) [sqm],
  - $k = 10$: Transportation network link length [km],
  - $k = 11$: Aggregated motivations connected with housing [trips],
  - $k = 12$: Aggregated motivations connected with work [trips],
  - $k = 13$: Aggregated motivations connected with other activities [trips],
  - $k = 14$: Aggregated motivations connected with freight traffic [trips];
- for *V2* the characteristics $A_{V2,Zj}(k)$ include energy consumption per unit area of the zone:

- $k = 1$: Passengers' cars [MJ/m^2],
- $k = 2$: Articulated heavy goods vehicles [MJ/m^2],
- $k = 3$: Heavy goods vehicles [MJ/m^2],
- $k = 4$: Small goods vehicles (vans) [MJ/m^2];

– for *V3* the characteristics $A_{V3,\,Zj}(k)$ include the quantity of postal infrastructure data in each zone, divided by the zone area:
  - $k = 1$: Post office [pcs./m^2],
  - $k = 2$: Post box [pcs./m^2],
  - $k = 3$: Parcel lockers [pcs./m^2].

The simulation and verification of zone matching were conducted for three scenarios of the courier parcel delivery system: **S1** – couriers deliver parcels directly to the customer's door, **S2** – couriers deliver parcels only to parcel lockers, from which they are picked up by customers, **S3** – deliveries are made to parcel lockers (as in **S2**), from where drones transport the parcels to the customer's door. Each parcel locker has a predefined number of drones, making this scenario an extension of S2 with an additional delivery stage.

In conducting the simulations and subsequent verification in the selected zones, the following assumptions were made for this stage of the study:

– nodes and links were introduced for the selected traffic analysis zone based on the division from the traffic model for the GZM Metropolis, the starting and returning node for the courier vehicle is chosen as the first node in the network,
– links have a uniform type and average speed, and their curvature is not considered,
– only residential buildings are included single-family houses and multi-family,
– parcel demand was determined based on the population of the zone, assuming **0,1** parcels per resident, demand is randomly distributed across specific locations,
– parcels have varying weights (randomly generated parcel structure), which affects the delivery feasibility by drones,
– the handling time (parcel drop-off) by the courier in **S1** follows a Poisson distribution: Pois(λ=120), while placing parcels in a locker follows Pois(λ=25),
– one courier is responsible for deliveries either to homes or parcel lockers, while in **S3**, each parcel locker is equipped with three drones that perform deliveries,
– the delivery location selection (by the courier and drones) is based on a greedy nearest-neighbor algorithm;
– simulation results – evaluation criteria for the *i*-th variant and selected zone (where $i = 0, 1, 2, 3$) are represented as $K(Vi)=[Tc,Dc,Ttd,Tdd,Dd]$ including:
  - *Tc* – courier delivery time [h],
  - *Dc* – distance traveled by the courier [km],
  - *Ttd* – total operating time of drones [h],
  - *Tdd* – drone delivery time (time required to complete all deliveries) [h],
  - *Dd* – total distance travelled by drones [km];
– based on the simulation results for variant *V0*, the estimated results for the remaining zones are derived using the following relationships for $i = 1, 2, 3$:
  - population size LM(*Vi*) where: $E1(Vi/V0) = \mathrm{LM}(Vi)/\mathrm{LM}(V0)$,
  - road connection length RC(*Vi*) where: $E2(Vi/V0) = \mathrm{RC}(Vi)/\mathrm{RC}(V0)$,

  - estimation of values for similar zones is performed using the following formulas: $K'(Vi) = E1(Vi/V0) \cdot K(V0)$ oraz $K''(Vi) = E2(Vi/V0) \cdot K(V0)$;
- the discrepancy analysis was conducted using the MAE (Mean Absolute Error) and MAPE (Mean Absolute Percentage Error).

3.2 Characteristics of variants – reference zone (base) and selection of similar zones

A zone for based variant (*V0*) for the experiment was selected to be the Podlesie district (*Z1* = 1084) on the outskirts of Katowice. The aim of selecting the base zone was to find a zone where delivering packages would be costly. It needed to be a zone with low development intensity, residential characteristics, and a low-class road network. Therefore, the zone has suburban characteristics, with single-family housing, low density, and a scarcity of functions beyond residential use.

The assessment of zone similarity was conducted based on the comparison of features describing each zone. Variants with different parameters were chosen to describe each zone based on the formula presented in 3.1 to assess the Euclidean distance between zones. Similarities in each variant were calculated, and the results are shown in Table 1.

Table 1. Assessment of similarities for all traffic zones

Zone	V1	V2	V3	Zone	V1	V2	V3	Zone	V1	V2	V3	Zone	V1	V2	V3
1001	1,608	40337	22,476	1023	0,815	25003	4,990	1045	0,126	1824	1,636	1067	0,228	5926	2,667
1002	1,032	15498	0,935	1024	0,361	7376	14,118	1046	0,230	7104	2,044	1068	0,441	1796	6,588
1003	1,027	7065	9,267	1025	0,316	653	2,614	1047	0,309	10927	2,614	1069	0,127	2722	2,080
1004	0,624	213508	4,319	1026	0,456	7628	14,101	1048	0,146	922	1,735	1070	0,227	21674	1,606
1005	0,525	16348	13,639	1027	0,405	10655	26,256	1049	0,123	1266	2,614	1071	0,133	3414	2,614
1006	0,582	6021	2,614	1028	0,148	11143	1,099	1050	0,148	69298	2,588	1072	0,064	669	1,518
1007	1,098	17128	5,471	1029	0,269	16301	1,119	1051	0,117	358	2,588	1073	0,111	5244	2,614
1008	0,534	51752	1,898	1030	0,234	5944	2,838	1052	0,135	6742	2,041	1074	0,073	977	2,614
1009	0,983	8104	0,909	1031	0,339	13857	1,957	1053	0,096	481	1,931	1075	0,051	5461	1,233
1010	1,146	8977	3,400	1032	0,237	3248	1,335	1054	0,148	5518	1,301	1076	0,085	56103	2,614
1011	0,465	6068	2,614	1033	0,555	4232	17,123	1055	0,144	11136	2,303	1077	0,099	13932	2,614
1012	0,441	7707	7,018	1034	0,367	21818	5,469	1056	0,183	18158	0,518	1078	0,105	143	2,614
1013	0,312	39189	0,538	1035	0,369	62920	3,280	1057	0,137	5589	2,614	1079	0,111	243	1,907
1014	0,475	63014	4,369	1036	0,121	475	2,794	1058	0,094	17620	2,614	1080	0,234	2141	5,552
1015	0,419	37292	1,638	1037	0,139	5723	2,614	1059	0,126	11026	2,614	1081	0,021	2045	0,686
1016	0,484	65938	1,028	1038	0,308	52809	2,614	1060	0,225	10225	3,206	1082	0,114	2534	0,854
1017	0,308	17159	2,429	1039	0,191	37702	5,404	1061	0,131	10416	2,614	1083	0,043	676	2,173
1018	0,460	69822	2,614	1040	0,473	29844	2,820	1062	0,131	3453	2,614	1084	0,000	0,000	0,000
1019	0,460	1964	2,614	1041	0,350	24785	1,079	1063	0,259	6278	2,369	1085	0,133	143	2,383
1020	0,187	58240	2,102	1042	0,165	6874	2,089	1064	0,288	10251	2,614	1086	0,118	8010	1,914
1021	0,213	22134	2,113	1043	0,119	30541	1,851	1065	0,128	3399	1,940	1087	0,138	1430	2,614
1022	0,215	16370	2,578	1044	0,173	1963	0,445	1066	0,128	1670	2,614	1088	0,138	2798	2,614

The zone for first variant (*V1*) was selected based on data related to urban development, extracted from the national GIS database (BDOT10k), and traffic motivation data from the local traffic model, based on traffic research. In the second variant (*V2*), zone was selected based on energy consumption data. It was obtained from the traffic model using the HBEFA database, which enables energy consumption. In this variant, fuel and electric vehicle consumption were aggregated together as energy consumption and recorded separately for each vehicle type present in the model. The third variant (*V3*) used data on the package infrastructure of the zone, where the number of post offices, post boxes, and parcel lockers was analyzed.

The map (Figure 2) shows locations of the one most similar zone in each variant. In variant *V1* and *V2* indicates the neighboring zones to the base zone, whereas *V3* indicates the residential area in the other part of the city.

The most similar zone for each variant has been compared across the other variants (Table 2), and its ranking position (besides the base zone) is shown. As can be seen each method of calculating zone similarity gives different results depending on the data used in the similarity analysis. For the next steps, the characteristics of the considered zones are presented in Table 3. This table also includes the linear value used for estimating the evaluation criteria of the distribution system.

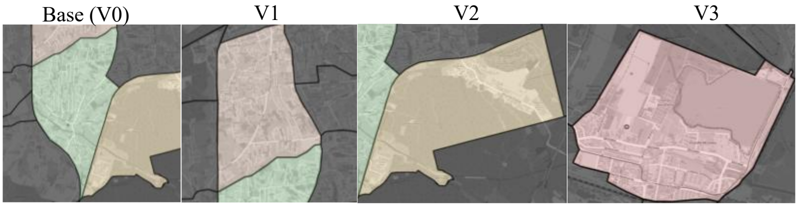

Fig. 2. Maps of the reference zone (V0 – base) and zones assessed as similar based on different characteristics for V1, V2, and V3.

Table 2. Similarity function values for zones V1, V2, and V3 and ranking position

Variant	Zone	**Similarity**			**Rank**		
		V1	*V2*	*V3*	*V1*	*V2*	*V3*
V1	1081	0,021	2045	0,686	1	19	4
V2	1085	0,1333	143	2,383	27	1	36
V3	1044	0,173	1963	0,445	39	17	1

Table 3. Characteristics of zones V0 - V3

	V0	*V1*	*V2*	*V3*
Number of nodes	82	55	10	63
Road connection length [km]	16,939	10,660	1,850	13,354
Number of single-family buildings	1547	1576	21	139
Number of multi-family buildings	74	181	62	291
Number of parcel lockers	7	8	1	4
Residential building area [m²]	194634	209960	21197	84360
Number of residents	4980	5369	2604	6537
Demand [pcs.]	498	536	260	653
E1(*Vi*/*V0*)	1	1,07811245	0,522891566	1,312650602
E2(*Vi*/*V0*)	1	0,629316961	0,10921542	0,788358227

3.3 Simulation study results, verification and approximation quality

For the selected zones, simulation studies were conducted, and results were estimated based on the first zone and similarity, following the presented methodology and adopted assumptions. Table 4 presents the simulation results for zone *V0* along with the estimation results – *K'* and *K"*. Table 5 shows the simulation results for the compared zones in variants *V1, V2*, and *V3*. Based on these two tables, the approximation was verified for the selected zones by assessing the error, which is presented in Table 6.

The conducted analyses indicate a clear advantage in estimating zone similarity based on a composite characteristic vector for *V1* and estimation based on the number of residents. This approach achieves an average absolute error of 12% for all criteria and simulated variants, which is an acceptable result. The second-best result was obtained using the approach based on postal infrastructure density (*V3*) and estimation based on road network density. In this case, MAPE is at 25%, which is at the threshold of acceptability. However, the approximation quality is more than twice as poor compared to the *V1* approach. It should be emphasized that the greatest discrepancies concern the distance traveled by couriers.

Table 4. Simulation study results for zone in *V0* and estimation for zones in *V1*, *V2,* and *V3*

Scenario		Criterion	***K(V0)***	*K'(V1)*	*K''(V1)*	*K'(V2)*	*K''(V2)*	*K'(V3)*	*K''(V3)*
S1		*Tc*	**16,824**	18,138	10,588	8,797	1,837	22,084	13,263
		Dc	**60,168**	64,868	37,865	31,461	6,571	78,980	47,434
S3	**S2**	*Tc*	**3,916**	4,222	2,465	2,048	0,428	5,141	3,087
		Dc	**5,987**	6,455	3,768	3,131	0,654	7,859	4,720
		Ttd	**12,976**	13,990	8,166	6,785	1,417	17,033	10,230
		Tdd	**4,524**	4,877	2,847	2,365	0,494	5,938	3,566
		Dd	**308,308**	332,390	194,023	161,211	33,672	404,700	243,057

Table 5. Simulation study results for the verification of results for zones in *V1*, *V2,* and *V3*

Scenario		Criterion	*K(V1)*	*K(V2)*	*K(V3)*
S1		*Tc*	17,787	3,108	12,878
		Dc	56,397	9,727	28,579
S3	**S2**	*Tc*	3,772	2,407	4,706
		Dc	12,017	4,220	5,127
		Ttd	14,105	2,973	9,617
		Tdd	4,751	3,056	5,402
		Dd	317,558	60,082	192,027

Table 6. Comparison of Estimated and Simulated Values

Scenario		*Criterion*	K(V1) to K'(V1)		K(V1) to K"(V1)		K(V2) to K'(V2)		K(V2) to K"(V2)		K(V3) to K'(V3)		K(V3) to K"(V3)	
			MAE	**MAPE**	**MAE**	**MAPE**	**MAE**	**MAPE**	**MAE**	**MAPE**	**MAE**	**MAPE**	**MAE**	**MAPE**
S1		*Tc*	0,351	1,975	7,199	40,475	5,689	183,049	1,271	40,880	9,206	71,487	0,385	2,993
		Dc	8,471	15,020	18,532	32,860	21,734	223,444	3,156	32,443	50,401	176,356	18,855	65,975
S3	**S2**	*Tc*	0,450	11,935	1,307	34,661	0,359	14,924	1,979	82,230	0,435	9,247	1,618	34,388
		Dc	5,562	46,288	8,249	68,647	1,089	25,817	3,566	84,506	2,732	53,289	0,407	7,937
		Ttd	0,115	0,816	5,939	42,104	3,812	128,228	1,556	52,330	7,416	77,117	0,613	6,374
		Tdd	0,126	2,650	1,904	40,081	0,690	22,594	2,562	83,832	0,536	9,918	1,836	33,985
		Dd	14,833	4,671	123,535	38,901	101,130	168,321	26,410	43,956	212,673	110,752	51,030	26,574
Error average			**4,273**	**11,908**	**23,809**	**42,533**	**19,215**	**109,482**	**5,786**	**60,025**	**40,486**	**72,595**	**10,678**	**25,461**

4 Conclusions and further research

The analysis of the obtained results confirms the validity of the proposed approach, which integrates micro- and macro-scale models through appropriately selected approximation functions and zone characteristics. To facilitate the implementation of this approach in real-world analyses and projects, further research is required to achieve a more comprehensive understanding of various factors influencing final outcomes and to identify additional dependencies. Given the complexity of the issue, the next stages of research will focus on the following areas:

- Analysis of factors affecting approximation accuracy, including the refinement of estimation and similarity functions and the development of new functions for assessing the compatibility of zone characteristics. Additionally, it will be necessary to examine how varying weights assigned to individual characteristics influence the accuracy of approximations. Furthermore, different demand estimation methods and the level of detail in transport network representation will be verified to determine their impact on the obtained results. Another key aspect will be analysing the relationship between the degree of zone similarity and approximation error.
- Sensitivity analysis, which should include the influence of errors in the characteristics of a given zone (parameters k) and the evaluation criteria (parameters K). This will picture the vulnerability of the particular data inaccuracy.
- Use of the empirical data to evaluate the model results. The survey data would be analysed to validate the model results. Additionally, data acquisition from the delivery companies should bridge the gap between model and operational data.
- Development of the simulation model, with further improvements focused on expanding the range of delivery system variants and refining the conditions for transport process execution. Planned enhancements include incorporating different starting points for couriers, introducing time windows for serviced units, and implementing advanced route optimization algorithms, such as LNS (Large Neighbourhood Search) or GA (Genetic Algorithm). Additionally, a broader scope of randomness will be introduced, especially regarding demand fluctuations, average vehicle speeds, and service times. This will enable the verification of the applicability of the proposed method in dynamically changing conditions.
- Additional analysis variants, including research on alternative delivery systems, such as the use of cargo bikes and integration with public transport. An essential direction will also be the inclusion of other demand segments, such as food distribution and retail supply chains. Furthermore, verifying the developed approach in different cities with varying transport and logistics conditions will allow for an assessment of the scalability and flexibility of the proposed method.
- Analysis of the impact of zone delineation on approximation feasibility, including an investigation of how different methods of urban area segmentation (including the number and size of defined zones) affect the accuracy of approximations and error magnitude. Evaluating various territorial segmentation methods, considering factors such as road network structure, traffic intensity, and urban development characteristics, will enable the formulation of consistent guidelines on the optimal delineation of zones for effective integration of micro- and macro-scale models.

Acknowledgement

This research was funded by the National Science Centre, Poland (project no. 2022/04/Y/ST8/00134). Project title: Energy optimal urban logistics as a service (E-

Laas). Project implemented as part of the call ERA-NET Cofund Urban Accessibility and Connectivity (EN-UAC China Call) organized by JPI Urban Europe and the National Natural Science Foundation of China (NSFC) funded from the European Union's Horizon 2020 research and innovation programme under grant agreement No 875022.

References

1. Bruzzone, F., Cavallaro, F., Nocera, S.: The integration of passenger and freight transport for first-last mile operations. Transport Policy 100, 31–48 (2021).
2. Cai, M., Hong, L., Xiong, C.: Data-driven traffic zone division in smart city: framework and technology. Sustainable Energy Technologies and Assessments 52, 102251 (2022).
3. Dong, H., Wu, M., Ding, X., Chu, L., Jia, L., Qin, Y., Zhou, X.: Traffic zone division based on big data from mobile phone base stations. Transportation Research Part C: Emerging Technologies 58, 278–291 (2015).
4. Gao, J., Zhen, L., Wang, S.: Multi-trucks-and-drones cooperative pickup and delivery problem. Transportation Research Part C: Emerging Technologies 157, 104407 (2023).
5. Jacyna, M., Żochowska, R., Sobota, A., Wasiak, M.: Decision support for choosing a scenario for organizing urban transport system with share of electric vehicles. Scientific Journal of Silesian University of Technology. Series Transport 117, 69–89 (2022).
6. Kim, K.: Identifying the structure of cities by clustering using a new similarity measure based on smart card data. IEEE Transactions on Intelligent Transportation Systems 21(5), 2002–2011 (2019).
7. Kłodawski, M., Jachimowski, R., Chamier-Gliszczyński, N.: Analysis of the Overhead Crane Energy Consumption Using Different Container Loading Strategies in Urban Logistics Hubs. Energies 17(5), 985 (2024).
8. Marujo, L.G., Blanco, E.E., Mota, D.O., Leite, J.M.L.G.: The use of public railway transportation network for urban intermodal logistics in congested city centres. Sustainable Rail Transport, 187–207 (2020).
9. McKenzie, G., Romm, D.: Measuring urban regional similarity through mobility signatures. Computers, Environment and Urban Systems 89, 101684 (2021).
10. Meng, Y., Li, S., Chen, K., Li, B., Zhang, J.E., Qing, G.: Research on spatiotemporal characteristics of urban crowd gathering based on Baidu heatmap. Archives of Transport 68(4), 41–54 (2023).
11. Plank, M., Lemardelé, C., Assmann, T., Zug, S.: Ready for robots? Assessment of autonomous delivery robot operative accessibility in German cities. Journal of Urban Mobility 2, 100036 (2022).
12. Qu, X., Zeng, Z., Wang, K., Wang, S.: Replacing urban trucks via ground–air cooperation. Communications in Transportation Research 2, 100080 (2022).
13. Wang, T., Lu, H., Sun, Z., Wang, J.: Towards higher efficiency and less consumption: control strategy and simulation for CAV platooning. Physica A: Statistical Mechanics and its Applications 613, 128518 (2023).
14. Wang, X., Zhen, L., Wang, S.: Optimizing an express delivery mode based on high-speed railway and crowd-couriers. Transport Policy 159, 157–177 (2024).
15. Wang, Y., Xiao, Y., Lai, J., Chen, Y.: An adaptive k nearest neighbour method for imputation of missing traffic data based on two similarity. Archives of Transport 54(2), 59–73 (2020).
16. Yang, B., Tian, Y., Wang, J., Hu, X., An, S.: How to improve urban transportation planning in big data era? A practice in the study of traffic analysis zone delineation. Transport Policy 127, 1–14 (2022).

17. Yang, G., Ma, H., Chen, K., Zhou, A.: Distribution Optimization for Connected Autonomous Vehicles (CAV) Considering Fuel Consumption Optimization. In: International Conference TRANSBALTICA: Transportation Science and Technology, pp. 205–222. Springer Nature Switzerland, Cham (2023).