Journal of civil engineering and transport **transEngin**

2026, Volume **8** Number **2** **ISSN 2658-1698, e-ISSN 2658-2120**

DOI: 10.24136/tren.2026.005

OPEN ACCESS

AN ALGORITHM FOR DETERMINING AREA-BASED AVAILABILITY OF PUBLIC TRANSPORT SERVICES USING DIGITAL TIMETABLES

Wojciech **Miechowicz** [1], Marcin **Kiciński** [2*], Piotr **Sawicki** [3]

[1] Poznań University of Technology, Faculty of Civil and Transport Engineering, Department of Transport Systems, Piotrowo 3, 61-158 Poznań, Poland, e-mail: wojciech.miechowicz@doctorate.put.poznan.pl, https://orcid.org/0000-0002-5243-9643

[2] Poznań University of Technology, Faculty of Civil and Transport Engineering, Department of Transport Systems, Piotrowo 3, 61-158 Poznań, Poland, e-mail: marcin.kicinski@put.poznan.pl, https://orcid.org/0000-0003-1843-3405

[3] Poznań University of Technology, Faculty of Civil and Transport Engineering, Department of Transport Systems, Piotrowo 3, 61-158 Poznań, Poland, e-mail: piotr.sawicki@put.poznan.pl, https://orcid.org/0000-0001-8392-8127

* Corresponding author

Reviewed positively: 15.01.2026

Information about quoting an article:
Miechowicz W., Kiciński M., Sawicki P. (2026). An algorithm for determining area-based availability of public transport services using digital timetables. *Journal of civil engineering and transport*. 8(2), 77-90, ISSN 2658-1698, e-ISSN 2658-2120, DOI: 10.24136/tren.2026.005

Abstract – This study develops a generic algorithm that determines the spatial availability of public transport services (PTS) expressed by the number of active trips within a defined time interval for a defined selected area. The research is conducted in the context of growing importance of spatial analysis in diagnosing transport exclusion and supporting sustainable mobility planning. The applied timetable data in GTFS format and a flexible spatial model allow the analysis of availability at two spatial levels: municipalities and cadastral units in Poland. The developed algorithm automatically assigns trips to territorial units, making the calculations easily repeatable and integrating timetable and spatial data within a unified analytical framework. The results indicate significant regional differentiation in the availability of PTS and distinct disparities between the municipal and cadastral levels. The algorithm also enables extended quantitative analyses, such as examining the relationship between the number of trips and population size, which demonstrates its broader analytical potential. The tool can be applied to identify areas with no public transport service (complete transport exclusion), monitor temporal changes, analyse demand–supply relationships, and support public transport network planning. The results confirm that the algorithm can support transport policy and decision-making aimed at more balanced public transport service availability.

Key words – algorithm, spatial analysis, GTFS data, public transport services, transport exclusion

JEL Classification – C18, C41, C55, C63, R58

INTRODUCTION

The absence of any active public transport line within a delimited area is an unequivocal manifestation of the exclusion of residents of that area from the possibility of meeting their mobility needs using public transport services – PTS. This phenomenon is commonly termed transport exclusion [1]. Where lines exist but their number or departure frequency is limited, there is a clear signal that such exclusion may occur. However, beyond the number of trips and, in essence, the number of available origin-destination connections, many other factors determine exclusion, including the duration of regular journeys, service reliability, walking distance to stops, monthly travel costs, and the degree to which the PTS offer is adapted to people with reduced mobility (Zmuda-Trzebiatowski et al. [1]). This paper focuses exclusively on the issue of determining the number of trips as an indicator of PTS availability. In public transport, a trip is a single scheduled run of a vehicle that follows one predefined sequence from its origin to its destination; it is the atomic unit of service supply. As indicated by Miechowicz et al. [2], the number of trips is an important factor influencing the perceived attractiveness of PTS.

Currently, many countries lack well-developed, standardised passenger-information systems. This limits access to data and increases the resources required to obtain and process such data (primarily gathering data from multiple sources). Timetable information is often dispersed, i.e., held by transport authorities, carriers or local-government units, and appears in heterogeneous forms: digital (both standardised and lacking digital standardisation) or analogue. This is particularly stark in Poland in regional bus transport, as demonstrated by Sawicki et al. [3] and Bar et al. [4]. This situation primarily leads to difficulties in the implementation of standardised passenger information formats, such as the General Transit Feed Specification (GTFS) [5] and the Network Timetable Exchange (NeTEx) [6], as well as in the comprehensive publication and dissemination of such information in the public domain, for example through a National Access Point (NAP). A complete, system-wide, publicly available, and regularly updated timetable implementation provides a fundamental source of planning information for both passengers and authorities. It underpins sound inferences about the current state of the PTS offer in each area and informs the shaping of transport policy.

The literature offers no direct treatments of algorithms for assessing the physical availability of public transport services, apart from the authors' earlier work (Miechowicz, Kiciński, Sawicki [7]). While some professional geospatial analysis tools allow for counting the number of services within predefined areas or polygons, such functionality is typically embedded in comprehensive, commercial GIS environments and intended for a wide range of advanced spatial analyses (e.g. ArcGIS by Esri) [8]. Existing GIS-based approaches, such as the Transit Frequency Methodology by UMN [9], rely on stop-based frequency measures implemented within comprehensive GIS environments. Several studies have used GTFS or GIS-based workflows to assess aspects of public transport systems, including service frequency and network accessibility. For example, Bárta [10] proposed a GIS methodology combining GTFS and Thiessen polygons to analyse transport supply across Hungarian cities, incorporating multiple quantitative indicators including frequency (e.g., proximity, density). Dardas et al. [11] developed a geospatial workflow that integrates near real-time GTFS data with GIS tools to assess transit system performance, while other research, e.g. Fan & Li [12], Fayyaz, Liu & Zhang [13], has used GTFS to evaluate spatial equity and accessibility measures of public transit.

A related strand addresses public transport accessibility, which has received more attention and is a multi-dimensional concept. The assessment of accessibility to PTS is treated in a dual manner. On the one hand, at the strategic level, accessibility is analysed across large territories, primarily for monitoring purposes from the perspective of an organising authority or a regulator. On the other hand, the problem is viewed operationally, from the perspective of an individual passenger planning a journey or the planning of a line by an authority or operator. In all these cases, static timetable information dominates, i.e., information without real-time vehicle locations.

Strategic assessments based on passenger information can be conducted in both spatial and temporal dimensions (e.g., Rosik et al. [14]). Where the number of services is used as an indicator, studies usually focus on large cities (e.g., Szłapacki [15]) or selected sub-regions (e.g., Beim et al. [16]). A characteristic feature of these studies is reliance on non-standardised, analogue timetables, particularly in regional bus transport. In highly urbanised areas with developed public transport systems, access to and use of digital information are much greater; an example is work on PTS accessibility in Szczecin (Stępniak et al. [17]). This trend is consistent with the growing interest in digital timetables at various levels of territorial administration (Miechowicz et al. [2]) and, as Bar et al. [4] show, is important to many stakeholder groups. As digitalisation progresses in the coming years, there is scope to use such information to monitor transport poverty (Rosik and Duma [18]).

Operational approaches are mainly reflected at the individual level, either for passengers assessing accessibility or for authorities or operators designing new routes. Mulley et al. [19] found that digital passenger information (e.g., via mobile apps) is particularly popular among regular users, whereas infrequent users more often rely on traditional channels such as operator websites, authority portals or call centres. The use of information varies by journey stage: planning, execution or disruption. Regardless of these differences, any journey-planning application must have current public-transport data in digital form alongside other functions (e.g., ticket purchase) to support mobility effectively (Solecka and Kiciński [20]). Ellison et al. [21] showed that higher-income individuals, who are less inclined to switch to public transport, are more likely to do so where ticket purchase is easy, which shifts the focus towards the behavioural domain of journey planning.

From the organiser's perspective, digital timetable data, especially GTFS, are used for multiple tasks, including modelling public-transport trips within the existing network under normal operation and during disruptions. Examples include optimisation of the bus network in Melbourne (Pemberton [22]) and maximising the level of service under a fixed investment budget in Charlotte while minimising the cost of new stops (Fan and Li [23]). Fayyaz, Liu, Zhang [13] proposed an algorithm to assess accessibility by public transport to jobs, education and services. Leveraging GTFS enabled analyses in five-minute intervals for each stop work that would have been far more labour-intensive using analogue data. Senegas et al. [24] used digital timetables to derive isochrones (15 and 30 minutes) and an interactive visualisation of accessibility for Lyon residents after the proposed launch of a new tramline. In extraordinary circumstances, e.g., floods, failures or other impediments, Ferretti et al. [25] used digital timetable information (stop locations and line alignments) to estimate passenger flows between selected areas of Singapore, revealing weak points in the city's PTS network.

Beyond static digital timetables, interest in real-time information, e.g. GTFS-RT, is also growing. This allows real-time vehicle visualisation and analysis of deviations from scheduled running times (Tóth et al. [26]), which is particularly relevant for monitoring public-service contracts between operators and authorities.

Based on the decision problem outlined and the literature review, the following methodological conclusions can be drawn:

- Although in the existing literature, the results of professional geospatial analysis tools are indeed used in studies of public transport availability and accessibility, no lightweight and explicitly defined algorithm has been proposed for assessing the physical availability of public transport services independently of comprehensive commercial GIS tools.
- In the literature, PTS availability is less frequently discussed than PTS accessibility. Instead, a physical availability as a part of accessibility is assessed at strategic and operational scales. In the former, it relies mainly on static, often non-standardised and even analogue timetable information, which hinders comparability of results and efficient procedures for assessing the current offer.
- In strategic-scale studies of PTS accessibility and availability, time-based measures and isochrones dominate, while approaches using a relatively simple, comparable supply indicator (the number of active trips) appear only sporadically and usually in the context of large cities.
- Progressing digitalisation and the availability of GTFS or NeTEx standards create conditions for replicable and scalable analyses based on a standardised data format.
- No studies were found that propose a strategic-scale accessibility assessment algorithm – one that counts active services within defined spatial units at adjustable resolutions and within specified time window.

In summary, the goal is *to introduce a generic algorithm to assess area-based public transport availability by counting the number of active trips within chosen time window and area, using static timetables of digital form*. The algorithm is validated using data from municipalities and cadastral units in Poland, and the results are interpreted in a transport-policy context.

1. ALGORITHM FOR ASSESSING PTS AVAILABILITY

KEY ASSUMPTIONS AND DATA SOURCES

The algorithm assumes the availability of standardised digital timetables. This study employs the most widely adopted standard worldwide – GTFS [5] – including in the Polish context. Other standards can also be used e.g., NeTEx, although they were beyond the scope of implementation. The GTFS standard is widely used mainly in passenger-information systems, such as journey planners, but also for data exchange between IT systems supporting, for example, spatial planning or transport modelling. The basic structure of the GTFS dataset includes the provider's name (organiser, operator or carrier), lines (routes.txt), services (trips.txt), the geometry of lines (shapes.txt where available), departure times from stops (stop_times.txt), stop geolocations (stops.txt), and service calendars (calendar.txt, calendar_dates.txt). In addition to the above-mentioned mandatory elements, the GTFS structure also contains optional elements, such as fares, attributes related to mobility limitations, and the spatial layout of interchange nodes. Digital timetables in GTFS format were obtained under the T-Included project (Sawicki et al. [3]) and cover the whole territory of Poland. In the research presented, the reference period is a weekday, 26 November 2024 (from 00:00 to 23:59).

Based on stop geolocations, a geospatial overlay algorithm was used to assign stops to the area polygons defining the areas that constitute the surface units. Poland's administrative division has three levels of administrative units: voivodeships (regions), counties and municipalities. For the purposes of the study, units constituting this lowest level i.e., municipalities were adopted. In addition, to obtain greater precision, analyses were also carried out for cadastral units, which are surface units used for land and building records. Both of these unit types, in addition to their boundaries defined by polygon vertices, are described by an identifier, a name and a number. Polygons for municipalities and cadastral units with names and identifiers were obtained from the portal of the Head Office of Geodesy and Cartography – HOGC [27]. These two types of units were used in the study; however, the algorithm can also use other polygons (e.g., hexagons), depending on needs and the availability of digital spatial data.

Additionally, the algorithm's count of services per municipality was compared with the municipality's population to test for correlation between these two variables, i.e., to test the null hypothesis of no correlation. Population counts for municipalities, as at 31 December 2024, were obtained from the resources of Statistics Poland [28].

Algorithm characteristics and procedure

This subsection presents the characteristics of the algorithm and the step-by-step procedure for determining the number of active trips within a defined area and time window. It is an extended version of the concept presented by Miechowicz, Kiciński and Sawicki [7]. For clarity, the workflow is illustrated in Figure 1 (schematic), while full pseudocode is provided in Annex A.

At the input, the algorithm receives a dataset in the form of structured GTFS files containing information, among others, on public transport services (trips), timetables and stops for a specified area. From the dataset, information on data providers (*agency.txt*) is first extracted – this can be, for example, a transport authority or a carrier. Lines (routes.txt) and the services associated with them (*trips.txt*) are assigned to a given data provider. Next, for each trip the algorithm checks in the calendar files (*calendar.txt, calendar_dates.txt*) whether it is operated on the reference day. If so, a status is stored for the trip by means of the parameter ref_day (equal to 1 when it is operated on the reference day, 0 otherwise). In the next step, for all trips for which the parameter *ref_day* = 1, timetables (*stop_times.txt*) are attached, and to the timetable entries information on stops (*stops.txt*) is assigned, including, among other, their geographic coordinates. Subsequently, for a given trip the stops are assigned using a geospatial overlay to the corresponding area polygon of the spatial unit. After completing all iterations for each data provider and for all trips, the number of unique trip identifiers falling within each unit is counted.

Results – Number of spatially distributed trips

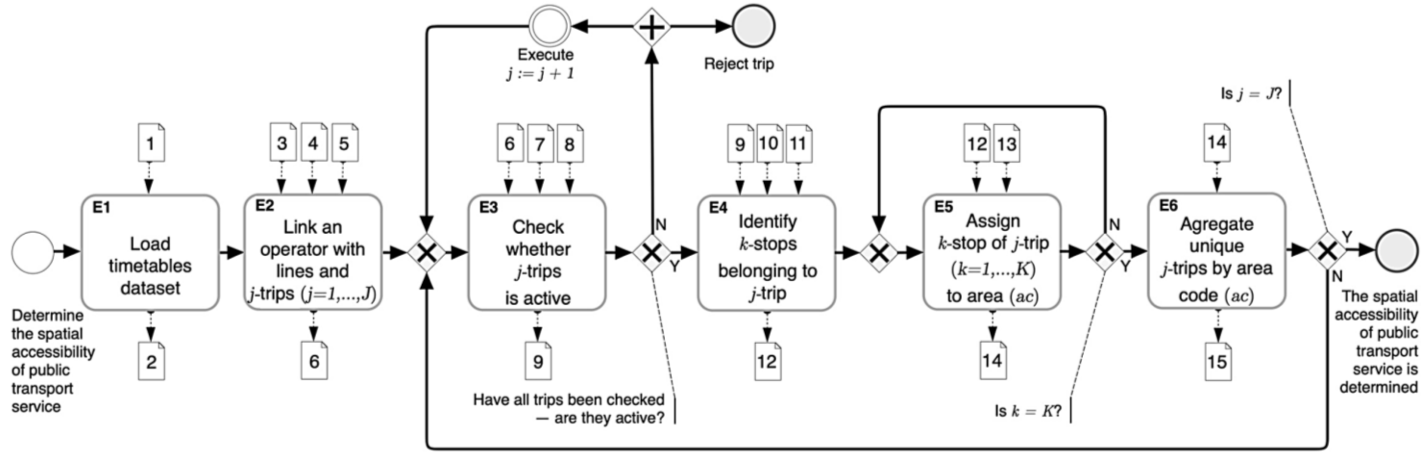
Tables 1 and 2 present example outputs obtained using the developed algorithm. Table 1 reports the number of trips in selected municipalities, while Table 2 gives the number of trips for cadastral units located within the spatial units included in Table 1.

Table 1. Overview of the number of trips identified within municipalities – a sample

id	file_name	agency_id*	agency_name**	munic_code	trips
1377	pkpic.rename.gtfs.zip	0	A	220701	2
1604	polregio.fixed1.gtfs.zip	0	B	220701	17
7341	GLE_ZKM_Lebork.fixed1.gtfs.zip	1	C	220801	181
7342	GLE_ZKM_Lebork.fixed1.gtfs.zip	1	C	220804	106
7343	GMB_MZK_MALBORK.redate.fixed1.gtfs.zip	1	D	220901	267
7344	GMB_MZK_MALBORK.redate.fixed1.gtfs.zip	1	D	220904	20
7613	zkm-gdynia.redate.gtfs.zip	6	E	221105	214
7623	zkm-gdynia.redate.gtfs.zip	8	E	221105	180
7628	zkm-gdynia.redate.gtfs.zip	9	E	221105	77
7636	zkm-gdynia.redate.gtfs.zip	17	E	221105	88
2113	wejherowo.gtfs.zip	2	F	221501	158
2114	wejherowo.gtfs.zip	2	F	221502	129
2115	wejherowo.gtfs.zip	2	F	221503	543
2116	wejherowo.gtfs.zip	2	F	221507	34
7619	zkm-gdynia.redate.gtfs.zip	7	E	221510	28
1389	pkpic.rename.gtfs.zip	0	A	221605	2
1640	polregio.fixed1.gtfs.zip	0	B	221605	36

* Note: *agency_id* corresponds to the original identifiers of organising authorities/operators/carriers in the source GTFS packages; therefore, it is not unique across the entire dataset, but it is unique within a GTFS package, see *agency name*

** A – PKP Intercity, B – PolRegio, C – ZKM LĘBORK, D – MZK MALBORK, E – ZKM Gdynia, F – MZK Wejherowo



LEGEND:
1 – Timetables in GTFS format; 2 – Successfully loaded GTFS feeds; 3 – File agency.txt (agency_id, …), 4 – File routes.txt (route_id, agency_id, …), 5 – File trips.txt (trip_id, route_id, service_id, …); 6 – List of j-trips (trip_id), $j = 1,..., J$; 7 – Reference day (rd); 8 – Files calendar.txt and calendar_dates.txt; 9 – Status of j-trip (trip_id);
10 – File stop_times.txt (trip_id, stop_id,…); 11 – File stops.txt (stop_id, stop_lat, stop_lon); 12 – Geolocation of k-stops belonging to the j-trip (trip_id); 13 – Polygons defining areas together with the area code (ac); 14 – Assignment of k-stop of j-trip to the area code (ac); 15 – Number of j-trips (trip_id) belonging to area code (ac).

Fig. 1. Algorithm for counting active trips within a delimited area and defined time window (based on [7])

Both tables show the record ID and the GTFS package name; the data-provider ID (organising authority, operator or carrier); the data-provider name; the municipality code in the case of municipalities, or the cadastral-unit code and its name in the case of cadastral units; and, in the last column, the number of unique trips crossing the given area.

Complete results for municipalities and cadastral units in the three selected voivodeships are shown graphically in Figure 2 and Figure 3. They contain maps of area-based PTS availability, with resolution at the level of municipalities and cadastral units, respectively. In both cases, an identical trip-count scale is used, allowing identification of areas with no provision (0), a minimal number of trips (1–10), then progressively increasing numbers (11–100; 101–500; 501–1000), and finally a very large number of trips (over 1000). A separate category comprises areas marked missing data, which reflects the absence of timetable information in any available form.

Table 2. Overview of the number of trips identified within cadastral units – a sample

id	file_name	agency _id*	agency _name **	cad-unit _code	cad-unit _name	trips
5947	pkpic.rename.gtfs.zip	0	A	220701_1.0010	10	2
6338	polregio.fixed1.gtfs.zip	0	B	220701_1.0010	10	17
30048	GLE_ZKM_Lebork.fixed1.gtfs.zip	1	C	220801_1.0001	Lębork obr 1	15
30049	GLE_ZKM_Lebork.fixed1.gtfs.zip	1	C	220801_1.0002	Lębork obr 2	15
30050	GLE_ZKM_Lebork.fixed1.gtfs.zip	1	C	220801_1.0003	Lębork obr 3	149
30051	GLE_ZKM_Lebork.fixed1.gtfs.zip	1	C	220801_1.0004	Lębork obr 4	83
30052	GLE_ZKM_Lebork.fixed1.gtfs.zip	1	C	220801_1.0005	Lębork obr 5	19
30053	GLE_ZKM_Lebork.fixed1.gtfs.zip	1	C	220801_1.0006	Lębork obr 6	104
30054	GLE_ZKM_Lebork.fixed1.gtfs.zip	1	C	220801_1.0007	Lębork obr 7	170
30061	GLE_ZKM_Lebork.fixed1.gtfs.zip	1	C	220804_2.0001	Chocielewko	5
30062	GLE_ZKM_Lebork.fixed1.gtfs.zip	1	C	220804_2.0007	Kębłowo Now.	12
30063	GLE_ZKM_Lebork.fixed1.gtfs.zip	1	C	220804_2.0011	Lubowidz	46
30066	GMB_MZK_MALBORK.redate.fixed1.gtfs.zip	1	D	220901_1.0002	2	8
30067	GMB_MZK_MALBORK.redate.fixed1.gtfs.zip	1	D	220901_1.0003	3	16
30068	GMB_MZK_MALBORK.redate.fixed1.gtfs.zip	1	D	220901_1.0005	5	43
30069	GMB_MZK_MALBORK.redate.fixed1.gtfs.zip	1	D	220901_1.0006	6	246
30070	GMB_MZK_MALBORK.redate.fixed1.gtfs.zip	1	D	220901_1.0008	8	192
32780	zkm-gdynia.redate.gtfs.zip	8	E	221105_2.0001	Pierwoszyno	1
32781	zkm-gdynia.redate.gtfs.zip	8	E	221105_2.0002	Mechelinki	28
32831	zkm-gdynia.redate.gtfs.zip	9	E	221105_2.0003	Rewa	77
32782	zkm-gdynia.redate.gtfs.zip	8	E	221105_2.0004	Kosakowo	105
32783	zkm-gdynia.redate.gtfs.zip	8	E	221105_2.0005	Kazimierz	49

*, ** see Table 1

A synthetic summary of the algorithm's outcome is presented in Table 3. It gives the percentage share of areas by bracketed number of trips, according to the analytical resolution – separately for municipalities and for cadastral units. The substantially higher share of areas with zero or a very low number of trips at the cadastral-unit level compared to the municipal level results primarily from the absence of public transport stops within the boundaries of many cadastral units. This effect reflects the fine spatial resolution of cadastral units rather than exclusively real deficits in public transport provision. At this level of detail, the results are more sensitive to polygon boundaries and to the spatial distribution of stops, which may lead to the identification of units without services even within municipalities that are generally well served. In some cases, a stop may be geographically located within a cadastral unit characterised by a low level of land use, while in practice it primarily serves a neighbouring unit with residential development, such as a housing estate.

Table 3. Share of delineated areas by number of trips

Trips in the time window [count]	Share of areas, in [%]					
	Warmian-Masurian		Pomeranian		Silesian	
	municips*	cad-units*	municips*	cad-units*	municips*	cad-units*
0 (none)	0.9	26.5	0.0	15.8	0.0	31.1
From 1 to 10	0.9	30.4	0.0	16.2	0.6	4.6
From 11 to 100	51.7	37.3	36.6	50.8	31.2	33.4
From 101 to 500	43.1	4.7	46.3	13.2	48.5	21.2
From 501 to 1000	1.7	0.9	12.2	2.2	6.5	4.2
Over 1000	1.7	0.2	4.9	1.8	12.0	4.5
Area with missing data	0.0	0.0	0.0	0.0	1.2	1.0

(*) municips – municipalities, cad-units – cadastral units

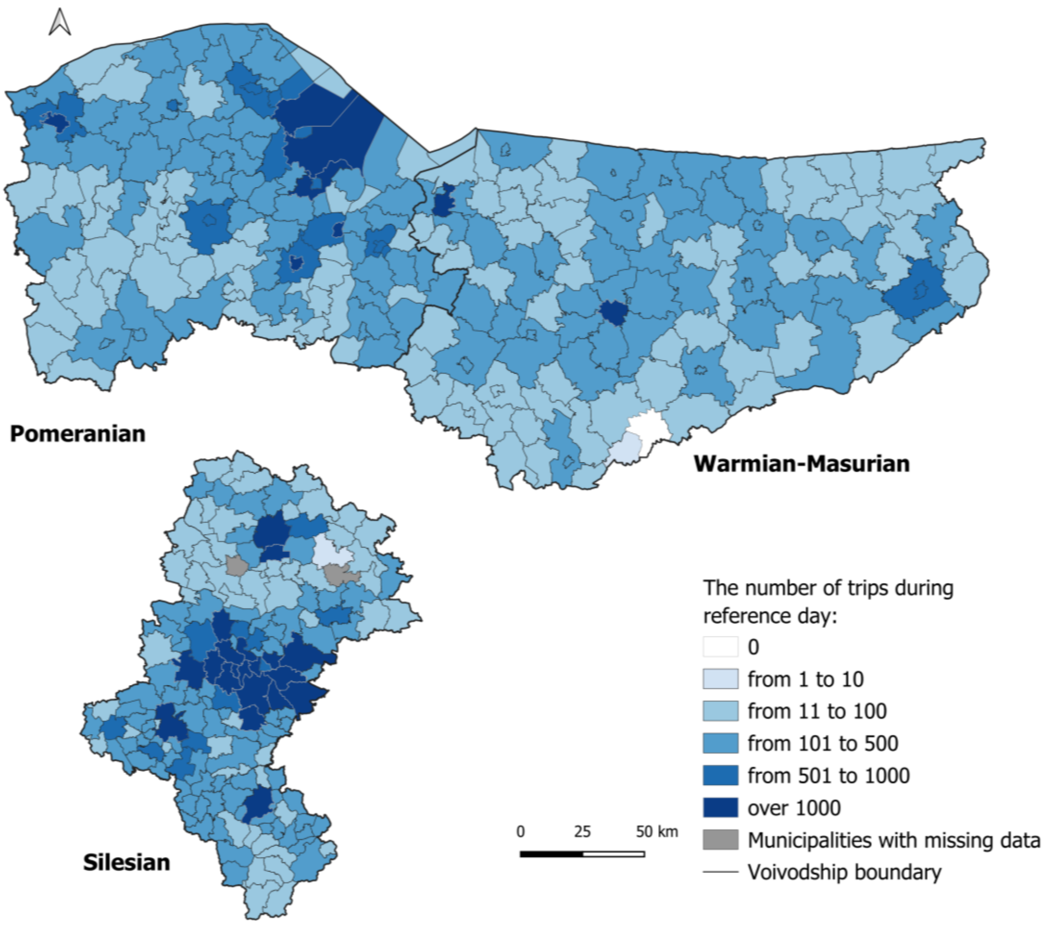


Fig. 2. Municipal availability of public transport services by number of trips during reference day – 26 Nov. 2024 (selected voivodeships)

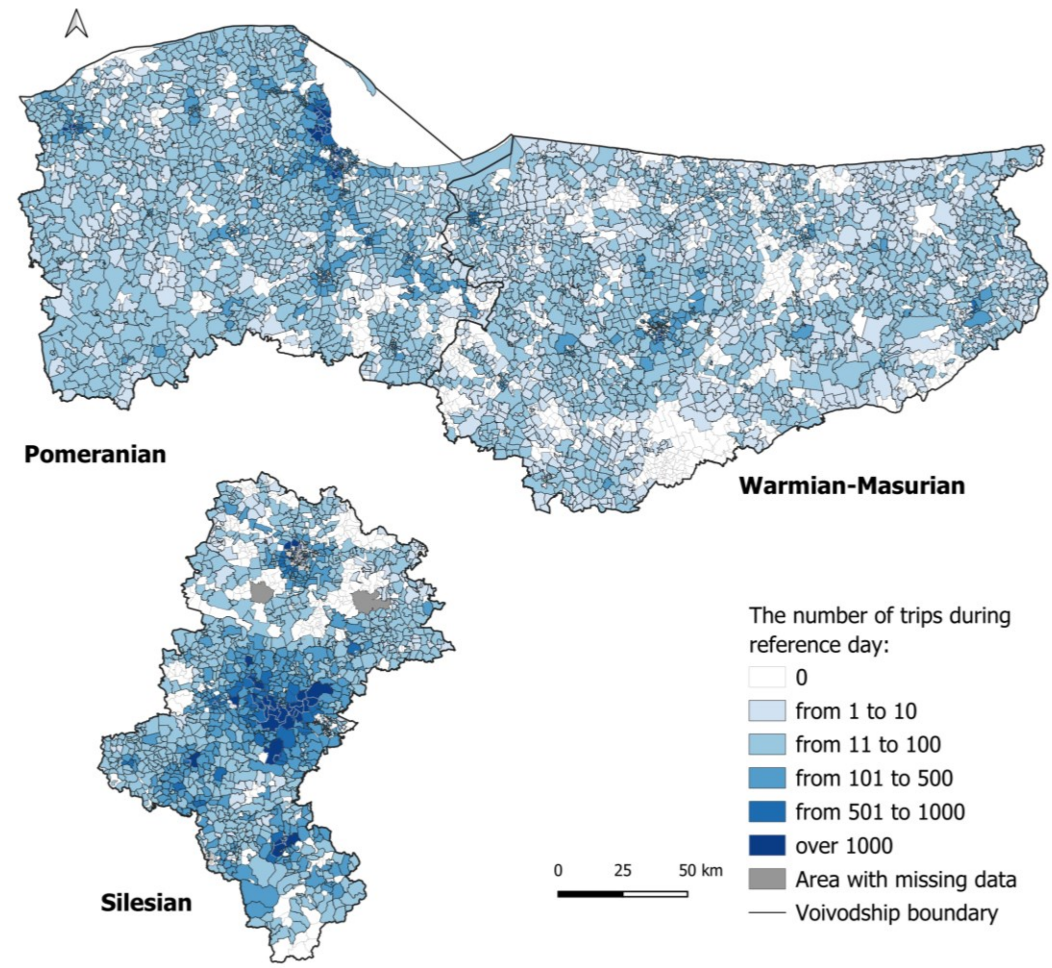


Fig. 3. Cadastral availability of public transport services by number of trips during reference day – 26 Nov. 2024 (selected voivodeships)

ASSESSMENT OF RELATIONSHIP BETWEEN THE NUMBER OF TRIPS AND THE POPULATION

In addition to the developed algorithm, for each municipality within the three voivodeships considered a correlation analysis was carried out between the number of trips and the number of inhabitants. The municipality was adopted as the unit of analysis, as current population figures are available and regularly updated at this level. For cadastral units such information is not available in open digital databases.

The results of the correlation analysis are presented graphically in Figure 4, and the corresponding indices are listed in Table 4. Because normality was not found for either variable (number of trips, number of inhabitants), Spearman's rank correlation coefficient R was used to determine the strength of association.

The results shown in Figure 4 and the computations summarised in Table 4 indicate a strong, and in the case of the Pomeranian voivodeship a very strong, correlation between the number of trips and the number of inhabitants across municipalities.

Table 4. Basic correlation indices for the number of trips and the number of inhabitants of municipalities (Spearman's R test)

Voivodeship	Number of municipalities N	Spearman's ρ*	p-value
Warmian-Masurian	116	0.685	0.000
Pomeranian	123	0.832	0.000
Silesian	165	0.730	0.000
Total	404	0.753	0.000

* Correlation coefficients are significant at $p < 0.050$

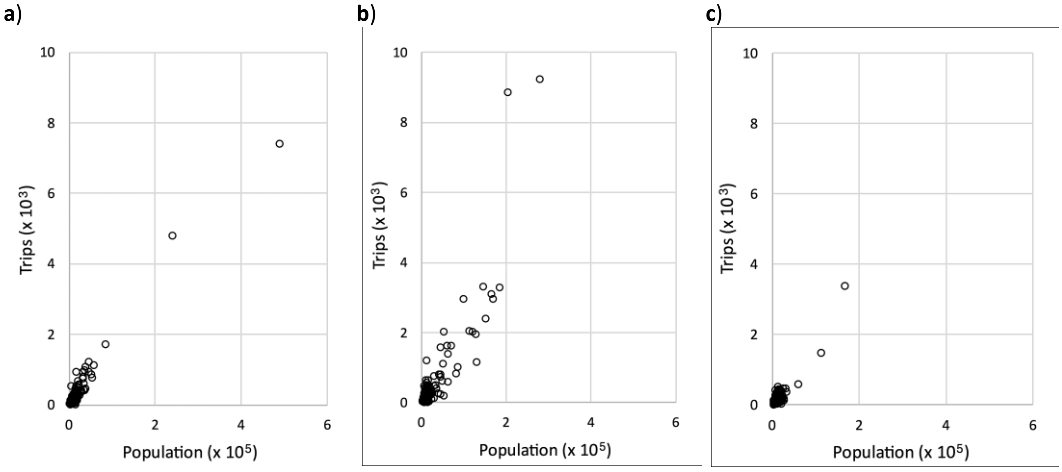


Fig. 4. Correlation between population and the number of trips at municipal level for the analysed voivodeships: (a) Warmian-Masurian, (b) Pomeranian, (c) Silesian. Spearman's ρ as in Table 4

2. Discussion of Results

PTS availability – differences by region and by scale

Figure 2 shows clear differences between the analysed voivodeships. Warmian-Masurian voivodeship is characterised by a predominance of municipalities with a low number of trips, which reflects a dispersed settlement pattern and low population density. In Pomeranian voivodeship, the situation is more balanced: the highest availability occurs within the Tri-City conurbation, whereas peripheral areas exhibit a moderate offer. Silesian voivodeship stands out with the highest trip intensity, especially across the Upper Silesian – Zagłębie Metropolis due to a dense agglomeration network; however, the northern and southern fringes of the voivodeship have a much poorer offer.

Analysis at the cadastral-unit scale (Figure 3) reveals much greater spatial variation. Within the same municipality, some areas are well served by public transport while others lack service entirely. This indicates that the fine-grained scale enables identification of local deficits that are invisible at municipal level. At municipal scale, the share of units excluded from transport (0 trips) and with minimal provision (1–10 trips) is small and amounts to 1.8% in Warmian-Masurian, 0% in Pomeranian and 0.6% in Silesian. At cadastral-unit scale, these values increase substantially: to 26.5% and 30.4% in Warmian-Masurian, 15.8% and 16.2% in Pomeranian, and 31.1% and 4.6% in Silesian, see Table 3. This shows that aggregate figures at municipal level mask both units that are deprived of an offer and those served at only a minimal level. The inverse relationship occurs for high and very high availability (≥501 trips). At municipal level, the share of such units is much higher than at cadastral-unit level, which reflects the concentration of intensive service in selected parts of large municipalities, especially county seats and conurbations, rather than uniform coverage across their entire area.

In summary, analysis at the cadastral-unit scale provides a picture closer to the reality experienced by passengers, as it reveals strong within-municipality differentiation in availability. At the same time, results at fine spatial resolution are sensitive to the exact course of unit (polygon) boundaries. This may, for example, lead to abundant provision being assigned to uninhabited land, while built-up areas with concentrations of residents lie just outside a given unit's boundary.

Relationship between the number of trips and the population

Correlation analysis at the municipal (see Fig. 4, and Table 4) indicates a strong relationship between the number of trips operated and the number of inhabitants. Spearman's rank-correlation coefficients reached values from 0.685 (Warmian-Masurian) to 0.832 (Pomeranian), which suggests that the transport offer is largely determined by population size. At the same time, this relationship does not fully explain spatial variation. In regions with dispersed settlement, such as Warmian-Masurian, municipalities with larger populations may still be weakly served by public transport. Conversely, in Silesian, municipalities with a small number of residents benefit from a dense agglomeration network. Due to the absence of publicly available population data for cadastral units, a correlation analysis could not be performed at that scale; however, the results in Table 3 indicate strong internal diversification of availability, independent of the average municipal population.

It should also be emphasised that the number-of-trips indicator alone does not provide a complete picture of the degree to which the offer is adapted to residents' needs. Among the available and analysed trips there are those intended for different functions, such as local school runs, inter-urban commuting or long-distance services stopping only in the largest localities or solely in the municipal seat. On the basis of the algorithm proposed, it is therefore not possible to infer unequivocally either the degree of transport exclusion or the attractiveness of the offer. A fuller assessment would require analysis of the ability to reach attractive destinations by means of the available trips (e.g., a town that is a centre of local services such as education, employment or healthcare).

Municipal versus cadastral-unit approaches

The results indicate the complementarity of the two approaches. The municipal scale is useful for macro-regional analyses, for comparison of administrative units and for assessing the relationship between the number of trips and population size. It provides a suitable tool for strategic planning and transport policy at voivodeship level. The cadastral-unit scale enables local diagnosis, reveals internal heterogeneity and makes it possible to identify areas actually at risk of an inadequate offer. It is particularly useful for designing intervention measures, such as launching feeder lines, improving stop availability or implementing targeted infrastructure investments. In a combined approach, the municipal scale provides the strategic background, and the cadastral-unit scale provides the basis for operational and socially salient actions. The algorithm used makes it possible not only to perform a quantitative assessment of availability but also to identify areas requiring intervention.

The analyses presented demonstrate that, in Poland, the problem of limited access to PTS is considerably more severe than suggested by statistics aggregated at municipal level. The true scale of the phenomenon becomes visible only at the more detailed, i.e. cadastral-unit, resolution. This clearly indicates the need to apply multi-scale analyses in the planning and coordination of public transport.

The proposed algorithm in the context of the state of the art

In prior studies assessing availability of public transport services, approaches based on analogue, unstructured timetables have predominated. Such studies have frequently been limited to large cities, have exhibited low repeatability and have offered limited spatial resolution. Current techniques that use digital timetable data are in large part employed to generate isochrones and to analyse selected time-based indicators. Extending the use of real-time timetable information in the GTFS-RT format requires considerably greater effort in data acquisition and subsequent processing.

Against this background, the algorithm proposed in this article takes explicit account of the digital format of PTS timetables and is formulated as a geospatial model, which enables the automatic assignment of trips to territorial units or to other user-defined polygons. The approach ensures full repeatability of calculations for defined time window of PTS operation. In addition, the results obtained make it possible to conduct extended quantitative analyses, for example the examination of the correlation between the number of trips and the number of inhabitants, which has hitherto not been feasible at this level of detail within a timeframe acceptable to decision-makers.

SUMMARY

MAIN RESULTS

The paper aimed to develop a generic algorithm for assessing area-based PTS availability, i.e., an algorithm for determining the number of active trips within a delimited area in a reference time interval. Key assumptions included the use of standardised digital timetables and the possibility of flexible definition of the boundaries of the area analysed. The algorithm was tested on the example of three voivodeships with markedly different population densities, using GTFS data and two spatial resolutions; two levels of area units were adopted for the analysis, i.e. municipalities and cadastral units.

The algorithm captures differences in PTS availability between spatial scales. It makes it possible to automatically assign the number of services to municipalities and cadastral units, which enables comparison of areas of different size and character. The algorithm also ensures repeatability of calculations and full comparability between higher-level administrative units, which is particularly important in regional analyses. Importantly, it enables a transition from dispersed data (stops, trips) to coherent indicators describing territorial units. Combined with other data, for example demographic, it allows extended relational analyses (e.g., relationships between number of services and population), which may constitute key performance indicators (KPIs) for monitoring the implementation of strategies or transport plans. Thus, the tool not only organises data but also increases their usefulness in decision-making processes by enabling the identification of both exclusion areas (no offer) and locations with graded concentrations of services, and by supporting different stages of research into the structure and performance of the PTS system. It should be noted that areas inherently unsuitable for PTS provision (e.g., due to the character of the analysed area) were not filtered; this should be the subject of an interpretive step.

From a transport-policy perspective, the algorithm can serve as a multi-level decision support tool. It helps identify areas at risk of transport exclusion and supports planning of targeted interventions (e.g., feeder lines or improved stop placement). It also allows for monitoring changes over time – such as seasonal shifts or the impact of policy updates – providing a consistent basis for temporal comparisons. By generating standardized indicators from fragmented timetable data, it fills gaps in existing transport statistics and enables cross-regional benchmarking. Because it works at multiple spatial scales, the tool goes beyond describing the current service offer and becomes part of a broader framework for assessing and balancing public transport availability, while helping to prevent transport exclusion.

FURTHER RESEARCH

Further research should extend the algorithm towards a comprehensive analytical decision tool. As a priority, it is justified to incorporate a qualitative aspect into the assessment of availability – so that not only the number of services but also their temporal, directional and functional structure are included. Such an approach would enable a more precise assessment of real PTS availability from the perspective of residents' needs.

A further direction is to integrate the algorithm with other data sources – e.g., employment data, income data or demographic structure. Combining these information layers will allow analysis of relationships between PTS availability and socio-economic factors. In this context, it is also worth developing analyses by mode (road versus rail) in order to better understand the functional and territorial determinants of the PTS system.

An important area of development remains the use of the algorithm in dynamic analyses and in forecasting modules to evaluate changes over time and simulate potential scenarios for network development. The inclusion of residents' mobility data (e.g., from GPS systems, mobile networks, automatic passenger-counting systems) could extend the scope of the assessment to aspects of actual demand.

In an application-oriented approach, further work should head towards developing a decision-support module that would support the planning and optimisation of public-transport services. Such a solution could be used to automatically indicate deficit areas and to assess the effects of changes introduced to the network structure. In summary, directions for further research should lead to the creation of a tool that would make it possible to conduct comprehensive analyses combining the supply of public-transport services with various dimensions of residents' mobility and to dynamically monitor changes over time.

ANNEX A

The pseudocode of the algorithm described in Section 1 is presented below.

Input: GTFS (*agency, stops, routes, trips, stop_times, calendar, calendar_dates*).*txt*, base map; sub-areas (polygons for areas and municipalities).

Algorithm steps:

1. Load GTFS files.
2. For each *agency_id* in *agency.txt* find all *route_id* in *routes.txt* and store *route_id*.
3. For each *route_id* find all *trip_id* in *trips.txt* and store *trip_id*.
4. For each *trip_id* find *service_id*.
5. Check in *calendar.txt* whether for the *service_id* found the following conditions for the reference day are satisfied, i.e,
   if conditions A-C are satisfied, i.e.,
   (A) the day falls within the dataset validity range;
   (B) *service_id* has value 1 for the weekday corresponding to the reference day;
   (C) in *calendar_dates.txt* there is no exception excluding the day (*exception_type* = 2),
   then set for the analysed trip the variable *ref_day* = 1 and proceed to Step 7,
   otherwise proceed to Step 6.
6. Check in *calendar_dates.txt* whether for the *service_id* found an addition for the reference day exists (for the relevant date *exception_type* = 1).
   if satisfied, set for the analysed trip *ref_day* = 1 and proceed to Step 7.
   otherwise proceed to searching the next trip in Step 4.
7. Check
   if the analysed *trip_id* runs on the reference day (*ref_day* = 1),
   then retrieve the timetable from *stop_times.txt* for the analysed *trip_id*,
   otherwise reject current *trip_id* and proceed to Step 4.
8. For the timetable found, find all *stop_id* and store them.
9. For each *stop_id* found, find in *stops.txt* the geographic coordinates of the stop (point).
10. Using geospatial overlay, assign each stop (point) to an area (polygon). Store the area code (*ac*).
11. After all iterations for all *agency_id*, count the number of unique *trip_id* in each area (i.e., trips per area).

Output: Number of trips per area.

ABBREVIATIONS

1. **GTFS** – General Transit Feed Specification,
2. **NAP** – National Access Point,
3. **NeTEx** – Network Timetable Exchange,
4. **PTS** – Public Transport Services.

ALGORYTM WYZNACZANIA OBSZAROWEGO WYSTĘPOWANIA USŁUG PUBLICZNEGO TRANSPORTU ZBIOROWEGO NA PODSTAWIE CYFROWYCH ROZKŁADÓW JAZDY

Celem pracy jest opracowanie generycznego algorytmu oceny obszarowego występowania usług publicznego transportu zbiorowego (PTZ), umożliwiającego wyznaczanie liczby aktywnych kursów w zdefiniowanym przedziale czasu dla skalowalnej rozdzielczości analiz. Badania są prowadzone w kontekście rosnącego znaczenia analiz przestrzennych w diagnozowaniu wykluczenia komunikacyjnego i planowaniu zrównoważonej mobilności. Zastosowane dane rozkładowe w formacie GTFS oraz elastyczny model przestrzenny pozwalają analizować dostępność w dwóch skalach: gminnej i obrębów ewidencyjnych. Opracowany algorytm automatycznie przypisuje kursy do jednostek terytorialnych, zapewnia powtarzalność obliczeń oraz umożliwia integrację danych rozkładowych i przestrzennych w jednym środowisku analitycznym. Wyniki wskazują na znaczące zróżnicowanie regionalne świadczenia usług PTZ oraz wyraźne różnice między poziomem gminnym a obrębowym. Algorytm może być również wykorzystywany w rozszerzonych analizach ilościowych, takich jak badanie zależności między liczbą kursów a liczbą mieszkańców, co stanowi przykład jego szerszego potencjału analitycznego. Narzędzie znajduje zastosowanie w monitorowaniu zmian w czasie, analizie relacji popytu i podaży oraz planowaniu rozwoju sieci transportowej. Wyniki potwierdzają jego przydatność jako uniwersalnego instrumentu wspierającego politykę transportową i procesy decyzyjne w zakresie równoważenia dostępności do usług PTZ.

Słowa kluczowe: algorytm, analiza przestrzenna, dane GTFS, publiczny transport zbiorowy, wykluczenie komunikacyjne.

AUTHOR CONTRIBUTIONS

– conceptualization, W.M. (Wojciech Miechowicz), M.K. (Marcin Kiciński) and P.S. (Piotr Sawicki);
– methodology, W.M., M.K. and P.S.;
– software, W.M.;
– validation, W.M. and M.K.;
– formal analysis, W.M., M.K. and P.S.;
– investigation, W.M., M.K. and P.S.;
– writing – original draft preparation, W.M., M.K. and P.S.;
– writing – review and editing, W.M., M.K. and P.S.;
– visualization, M.K. and P.S.;
– supervision, P.S.;
– project administration, M.K.

All authors have read and agreed to the published version of the manuscript.

REFERENCES

[1] Zmuda-Trzebiatowski P., et al. (2025). Analiza skali wykluczenia od transportu zbiorowego w województwie zachodniopomorskim [In:] Rychlewski J., Fierek S. (red.): *Modal Shift for Sustainable Mobility*. XV Poznań Scientific and Technical Conference 2025, 219-222.

[2] Miechowicz W., Kiciński M., Miechowicz I., Merkisz-Guranowska, A. (2024). The Attractiveness of Regional Transport as a Direction for Improving Transport Energy Efficiency. *Energies*, 17(19), 4844. https://doi.org/10.3390/en17194844.

[3] Sawicki P., et al. (2025). *Analiza skali wykluczenia komunikacyjnego na obszarze Polski wraz z rekomendacjami zmian legislacyjnych w kontekście publicznego transportu zbiorowego*. Raport końcowy, cz. A, 4/2025. GOSPOSTRATEG-V/0005/2021, Poznań, Gdańsk, Katowice, Warszawa, 74 s. (unpublished report).

[4] Bar K., et al. (2023). Cyfryzacja informacji pasażerskiej w Polsce – przyczyny niepowodzeń. *Biuletyn Komunikacji Miejskiej*, 168, 88-93.

[5] Mobility Data, *General Transit Feed Specification*, https://gtfs.org (accessed: 05 Sep. 2025).

[6] European Committee for Standardization – CEN. *Network Timetable Exchange*. PD CEN/TS 16614-(1-3): 2014.

[7] Miechowicz W., Kiciński M, Sawicki P. (2025). Assessment of the accessibility of regular public transport services in a regional context. [In:] Rychlewski J., Fierek S., *Modal Shift for Sustainable Mobility*. XV Poznań Scientific and Technical Conference, 215-218.

[8] ESRI. *ArcGIS Pro 3.6 - Calculate Transit Service Frequency (Public Transit).* https://pro.arcgis.com/en/pro-app/latest/–tool-reference/public-transit/calculate-transit-service-frequency.htm (access date: 27/01/2026).

[9] University of Minnesota Accessibility Observatory. (2015). *Transit service frequency methodology*. https://maps.umn.edu/transit-frequency/ (access date: 27/01/2026).

[10] Bárta, G. (2022). GIS-based methodology to analyse the public transport supply: Hungarian case studies. *Transportation Research Procedia*, 63, 2279–2288. https://doi.org/10.1016/j.trpro.2022.06.263.

[11] Dardas A., et al. (2022). A geospatial workflow for the assessment of public transit system performance using near real-time data. *ISPRS International Journal of Geo-Information*, 11(12), 628. https://doi.org/10.3390/ijgi11120628.

[12] Fan W., Li Y. (2019). *Optimizing transit equity and accessibility by integrating relevant GTFS data performance metrics*. Final Report, Center for Advanced Multimodal Mobility Solutions and Education, University of North Carolina at Charlotte. https://rosap.ntl.bts.gov/view/dot/57063

[13] Fayyaz S.K., Liu X.C., Zhang G. (2017). An efficient GTFS-enabled algorithm for dynamic transit accessibility analysis. *PLoS ONE*, 12(10), e0185333. https://doi.org/10.1371/journal.pone.0185333

[14] Rosik P., et al. (2017). *Multi–modalna dostępność transportem publicznym gmin w Polsce*. IGiPZ PAN, Warszawa.

[15] Szłapacki W. (2024). Dostępność transportu publicznego w wybranych średniej wielkości miastach Pomorza Środkowego. *Rozwój Region. i Polit. Region*., 71, 31–48.

[16] Beim M., et al. (2019). *Badania dostępności publicznego transportu zbiorowego w podregionie pilskim*. PKGK, 22 (4), 95–118.

[17] Stępniak M., et al. (2019). The impact of temporal resolution on public transport accessibility measurement: Review and case study in Poland. *Journal of Transport Geography*, 75, 8-24. https://doi.org/10.1016/j.jtrangeo.2019.01.007.

[18] Rosik P., Duma P. (2025). Transport poverty in Poland – a spatial approach to public transport demand and supply. *Geographia Polonica*, 98(3), 255-276. https://doi.org/10.7163/GPol.0302.

[19] Mulley C., Clifton G. T., Balbontin C., Ma L. (2017) Information for travelling: Awareness and usage of the various sources of information available to public transport users in NSW. *Transportation Research Part A: Policy and Practice*, 101 (7), 111–132. https://doi.org/10.1016/j.tra.2017.05.007.

[20] Solecka K., Kiciński M. (2022) A Multi-Criteria Evaluation of Applications Supporting Public Transport Users. *Energies*, 15(10), 3493-1-3493-25. https://doi.org/10.3390/en15103493.

[21] Ellison R.B., Ellison A.B., Greaves S.P., Sampaio B. (2017) Electronic ticketing systems as a mechanism for travel be–haviour change? Evidence from Sydney's Opal card. *Transportation Research Part A: Policy and Practice*, 99(5), 80-93. https://doi.org/10.1016/j.tra.2017.03.004.

[22] Pemberton S. (2020). Optimising Melbourne's bus routes for real-life travel patterns. Case Studies on *Transport Policy*, 8(3), 1038-1052. https://doi.org/10.1016/j.cstp.2020.04.002.

[23] Fan W., Li Y. (2019). *Optimizing Transit Equity and Accessibility by Integrating Relevant GTFS Data Performance Metrics*. Final Report: Project 02, Center for Advanced Multimodal Mobility Solutions and Education / The University of North Carolina at Charlotte. https://rosap.ntl.bts.gov/view/dot/57063.

[24] Senegas S., Furno A., Palmas P., Simon J. (2023). Framework for the Analysis and Enhancement of the Accessibility of Large-Scale Urban Transit Networks: A Data-Driven Study for the City of Lyon, France. *Transportation Research Record*, 2678(2), 35-47. https://doi.org/10.1177/03611981231172939.

[25] Ferretti M., et al. (2018) Weak Nodes Detection in Urban Transport Systems: Planning for Resilience in Singapore, *2018 IEEE 5th International Conference on Data Science and Advanced Analytics* (DSAA), Turin, Italy, pp. 472-480. https://doi.org/10.1109/DSAA.2018.00061.

[26] Tóth R., Ispány M., Zichar M. (2025). Interactive Visualization for the GTFS and GTFS-RT Data of Budapest. *ISPRS International Journal of Geo-Information*, 14(7), 245. https://doi.org/10.3390/ijgi14070245.

[27] Head Office of Geodesy and Cartography (2025) https://mapy.geoportal.gov.pl: A03 Municipalities, A06 Cadastral units (in Polish).

[28] Statistics Poland (2025). *Population. Size and structure and vital statistics in Poland by territorial division in 2024*. As of 31 December, https://stat.gov.pl (in Polish).