Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 42 (2020) 100–108

Transportation Research Procedia

www.elsevier.com/locate/procedia

46th European Transport Conference 2018, ETC 2018, 10-12 October 2018, Dublin, Ireland

Benchmarking public transport level-of-service using open data

Tuomas Palonen[a,*], Riku Viri[b]

[a]Ramboll Finland, Tampere, Finland
[b]Transport Research Centre Verne, Tampere University, Tampere, Finland

Abstract

More efficient public transport is needed as urbanisation continues and cities grow. To identify best practices, there is a need for simple methods to benchmark public transport in different cities. This study suggests a simple method for benchmarking public transport level of service with available open data. The presented method provides data on the number of inhabitants reached by certain headway classes in defined urban areas, based on approximated acceptable walking distance to the nearest stop. Given that the definition of urban area is harmonized, the method can quickly provide comparable public transport level of service data between cities.

Keywords: public transport; benchmark; level of service; headway; open data; geodata; GTFS

1. Introduction

Urbanisation is a continuing global megatrend. Demand for urban transport grows together with urban population; the number of trips and average time spent in traffic daily keeps unchanged despite changing lifestyles. Traffic problems cannot be solved with digitalisation and automatisation innovations alone. Economical, ecological and effective urban transport system should include a robust public transport system.

Each city has its own means of organising and providing public transport. Partially as a result, the public transport level of service (and its definition) varies between cities. There are no established methods to benchmark

* Corresponding author:
E-mail address: tuomas.palonen@ramboll.fi

10.1016/j.trpro.2019.12.010

the public transport level of service. To discover which cities have the best practices in public transport, a level of service benchmarking method or tool is required.

The level of service in public transport is a vague concept. Public transport customers, organisers and operators all have slightly different ideas on what the level of service stands for. The definition of level of service can be split in qualitative and quantitative features. Qualitative features are based on opinions or experiences and they cannot be measured directly, such as ride comfort, ease-of-use or feeling of safety. These features are typically measured with passenger surveys. Quantitative features can be measured directly, such as frequency, operating span, ride time or walking distance to a stop.

Comparing and analysing qualitative features between cities requires complex and coordinated work and is time-consuming but it would provide in-depth and extensive results. Quantitative features can be measured and analysed relatively fast, but the results yield only an outline on the subject.

The method presented here concentrates on fast process and results, so only selected quantitative level of service features are measured. Firstly, the method presented here is designed for fixed route and fixed schedule public transport. The method is based on the assumption that public transport frequency and walking distance to a stop are the determinative components to public transport level of service. The significance of frequency is emphasized by Walker (2012) and Mulley et al. (2018), among others. Another similar benchmark method suggested by Poelman and Dijkstra (2015) also regards frequency and walking distance as fundamental components to level of service.

2. Methods and Data

The presented public transport level of service benchmarking method requires population data grid, land use data and public transport timetable data as inputs. In addition, a boundary defining the geographical area is required. In European context, these data are typically open and publicly available.

Poelman and Dijkstra (2015) have described a similar method to benchmark the access to public transport. They suggest that in order to benchmark public transport in cities, the geographic definition of a city has to be harmonized and the population distribution needs to be taken into account. Poelman and Dijkstra (2015) calculated walking distances to stops, which yields highly accurate stop coverage areas. This calculation requires street network data and an additional calculation step, which is not required in the method presented here.

In brief, in this method, the population from the population grid is first disaggregated to land use with residential features. This is done because the land use data typically has higher resolution than the population grid data. Next, the coverage of different headway classes of public transport is calculated in relation to population in built-up urban areas.

2.1. Data sources

European Union statistical office Eurostat provides a 1 km² population distribution grid (EFGS 2011). The latest grid dataset is based on the 2011 census. European-wide censuses are composed every ten years; next census is planned for 2021. More recent and detailed population data could be used if such data is available. The 2011 grid contains 1 953 286 square kilometre cells and it covers the EU-28 and EFTA countries[1].

Pan-European land use open data datasets are available. Corine Land Cover (CLC) 2012 update covers all the land use in 39 European countries while more detailed Urban Atlas (also updated 2012) covers 693 functional urban areas (FUA) within Europe.

CLC is a European-wide open dataset on land use. Land cover data contains 44 classes that describe the land use across Europe. Population is mainly distributed on two land use classes numbered 111 and 112. These classes include all areas that contain residential features and cover at least 30 % of land surface. Minimum mapped area size in CLC is 25 hectares.

[1] EU countries included in population distribution grid are Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden and United Kingdom. Iceland, Liechtenstein, Norway and Switzerland are EFTA countries.

Urban Atlas (UA), a high resolution land use dataset derived from CLC is also available as open data. UA does not cover the same geographical extent as CLC but provides detailed land use data on functional urban areas. In general, UA includes cities with more than 100 000 inhabitants. UA has six land use classes that contain residential features. The classification is based on the amount of artificial surface, ranging from continuous urban fabric covering at least 80% of the surface to isolated structures. The concerned class numbers are 11100, 11210, 11220, 11230, 11240 and 11300. UA is significantly more detailed than CLC; the minimum mapped area size is 0.25 hectares.

For public transport departures, this benchmark uses open GTFS data. GTFS (General Transit Feed Specification) is a global specification used to publish public transport routes and timetables for journey planning purposes, originally created by Google (Google 2018a). Since many journey planning applications rely on GTFS data availability, several cities and major operators provide their route and timetable information in this format. Therefore, the use of GTFS in the benchmark makes it possible to have same kind of workflow with most of the benchmarked cities[2].

2.2. Calculating departures for stops

GTFS data is provided in multiple text files. These files consist of e.g. information about routes, stop locations, stop departure times, agencies, services and calendar data and exceptions of days the services are driven (Google 2018b).

First, a lookup date is defined[3] to filter the results only to show departures on that given day. For this benchmark, the day selected is a normal workday Tuesday. Since the GTFS data usually has timetable data for longer periods, the workflow then only filters to the list of services that are driven on this day based on the calendar and calendar exceptions data. The services then link to the actual trip data through routes date, so at this point, it is known which trips are driven on the given day.

Every stop along the routes of different transport modes (e.g. bus or tram) has an arrival and departure time for every trip. The last stop of route has a timestamp of arrival and departure, as also does the first stop of the next route. Therefore, the last stops of every trip needs to be filtered out. This is done by using the stop sequence numbering on the data as no passenger can board the vehicle on its last stop (since they board the first stop of the next trip). If both the last and first stop are counted, it creates duplicate departures on the final stops, effectively doubling their frequency when mapping the data.

A list of stops with the count of departures for the given day is calculated. This data can then be filtered to only show the departures of a desired timeframe; 6:00 to 22:00 is used on the calculation. The results can be further filtered to show departure count for different transport modes or different transport agencies operating in the area, if needed.

Last, these stops are matched to their location by using the stop data of GTFS. At this point, every stop has geo-coordinates and departure count on desired filter level. The data is then exported as spreadsheet to be used in next workflow, i.e. mapping the level of service.

2.3. Mapping the level of service

In order to do a viable benchmark, the benchmarked cities have to have harmonized definitions and borders. OECD has developed a harmonized definition of functional urban areas in collaboration with the EU. Typically, a functional urban area features a city and the commuting zone around it (OECD 2013). This definition is suitable for benchmarking urban public transport.

Next the areas that have predominant residential use are selected from the land use dataset and clipped with the FUA borders. In CLC dataset, the relevant land use classes are 111 and 112. In UA dataset relevant classes are

[2] Most of the cities and operators have their public transport data available in GTFS format. However, for example United Kingdom uses its own TransXChange data standard that is incompatible with GTFS schema.
[3] All the public transport timetable data used in this study is dated 7.8.2018.

11100 to 11300. If available, UA dataset should be used, since CLC dataset is more inaccurate and might contain inconsistencies between countries[4]. Typically, UA dataset covers cities of at least 100 000 inhabitants. Fig. 1 shows UA land use with overlaid population distribution grid in city of Tampere, Finland for reference.

Population distribution grid is next clipped with the residential land use to distribute the population from the grid to appropriate land use areas. Here the population from each grid cell is distributed homogenously within residential land use.

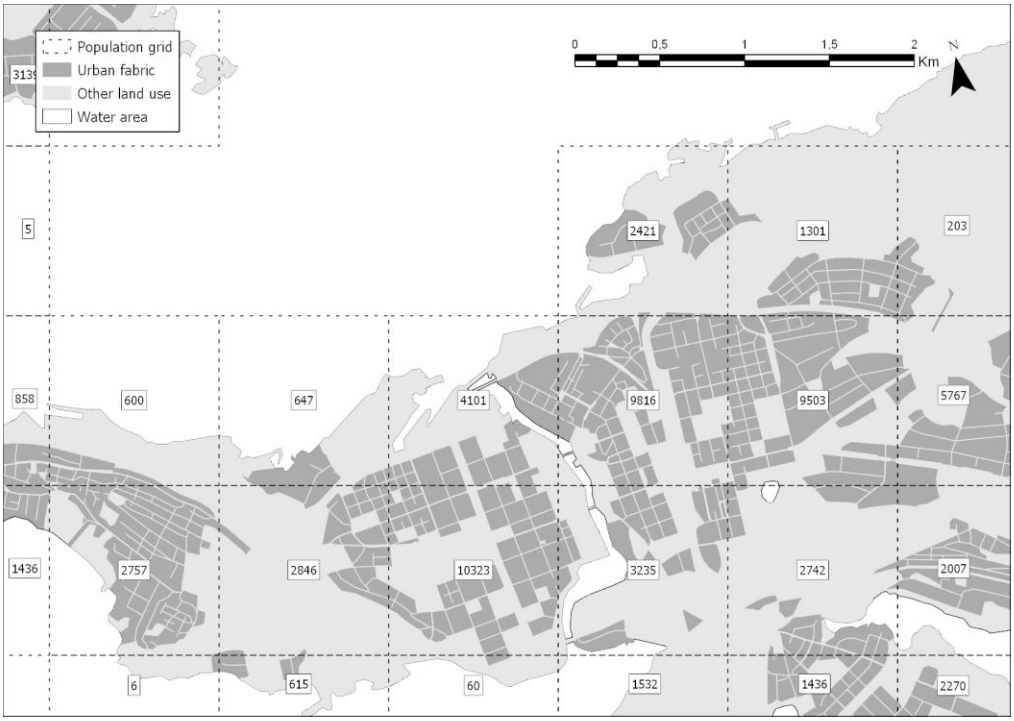

Fig. 1 Dark grey areas represent residential land use in Tampere, Finland. A population grid is overlaid on the map; the number in each grid cell represents the cell population. Using geoprocessing, the population is assigned to the dark grey areas within each cell.

Next the public transport stops are placed on map based on their coordinates. Each stop is given a buffer, representing the catchment area (Fig. 2). Finnish Transport Agency (2015, pp. 26) considers 400 metres the maximum walking distance to a stop of highest level of service. Transit Cooperative Research Program (TCRP 2003, pp. 3-10) has also concluded that most passengers are willing to walk at most a quarter-mile (or about 400 metres) to a public transport stop, and even farther to rail transport stops.

Since this method does not define stop catchment areas using actual walking distances on the streets and road network but circular buffers, a straight line conversion has to be made. Finnish Transport Agency (2015, pp. 28) suggests that a distance along the network is 1.3 times the straight distance. Therefore, stop buffer radius is set to 300 metres.

[4] CLC data is composed nationally and the data building processes may vary by country. For example Finnish CLC data does not include class 111; some urban areas that should be marked 111 are marked 121 instead.

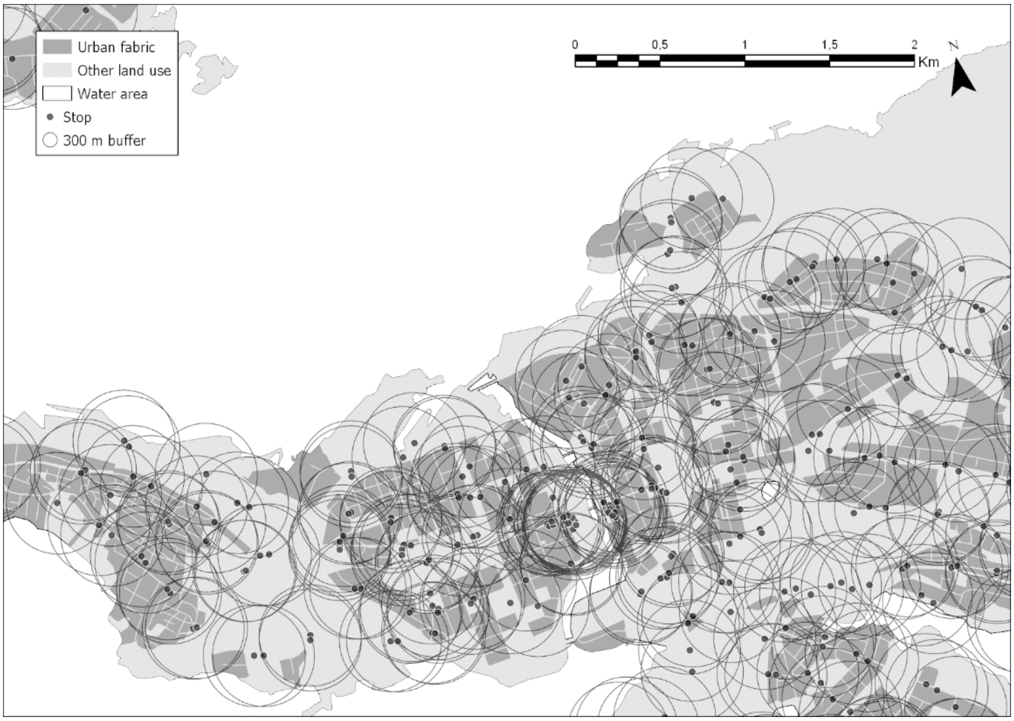

Fig. 2 Catchment areas (black circles) are appointed to the public transport stops (black dots) with a 300 metre buffer.

Next, stops are classified to eight classes depending on the number of departures within selected time period. Here a span from 6:00 to 22:00 is used; the length of the span is 16 hours. The classification presented in Table 1 is based on headway; the shorter the headway, the better the level of service.

Table 1. Headway classification.

Frequency; departures in total per 16 hours	Frequency; departures per hour on average	Average headway, in minutes
> 192	> 12	< 5
192–96	12–6	5–10
96–64	6–4	10–15
64–48	4–3	15–20
48–32	3–2	20–30
32–16	2–1	30–60
16–8	1–0.5	60–120
< 8	< 0.5	> 120

Finally each stop class catchment areas are clipped with inhabited areas, as shown in Fig. 3. The public transport headways can be seen at a glance on the map and also possible gaps in public transport service can be noticed.

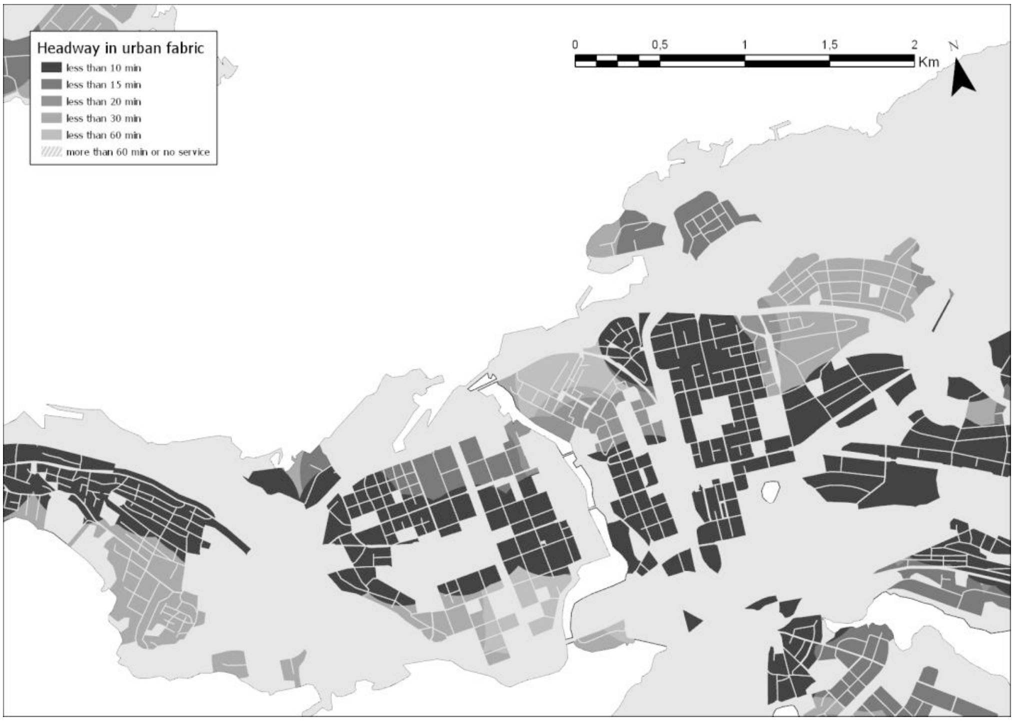

Fig. 3 Public transport service headway on areas with residential function. White areas are not within 300 metres (direct distance) to a public transport stop. The extent of this figure does not contain areas with two of the lowest classes of headway.

The inhabited area covered by buffered stops is compared to the original area of inhabited areas, which yields a percentage. This percentage is used to calculate the population within the catchment area of stops. Results now represent the extent and amount of inhabitants that are within the catchment area of each public transport headway class. The data can be now presented on numeric format.

3. Results

Given that appropriate data is available, using this method yields an approximate number of inhabitants within the reach of different public transport headway classes. These numbers can be used to benchmark the level of service of public transport in different cities.

In Fig. 4 there are 11 different cities that have been analysed using the method presented here[5]. The cities are ranked by the relative amount of population within the reach of the headway of 5 minutes or less. In Zürich, Helsinki and Bern about one third of the population is within the reach of headway of 5 minutes or less. In Manchester, however, public transport reaches relatively more inhabitants than the other cities, but only one out of ten inhabitants are within the reach of headway of 5 minutes or less.

[5] The used public transport data sources are:
Berlin: VBB (2018) VBB-Fahrplandaten via GTFS, Available at: https://daten.berlin.de/datensaetze/vbb-fahrplandaten-gtfs
Bern & Zürich: geOps (2018) Public Transportation Feed for Switzerland [converted from fahrplanfelder.ch], Available at https://gtfs.geops.ch/
Copenhagen: Transitfeeds (2018) Denmark GTFS, Data originally by Rejseplanen [downloaded from https://transitfeeds.com/p/rejseplanen/705]
Helsinki: HSL (2018) Open data, Available at https://www.hsl.fi/en/opendata
Jyväskylä: Jyväskylän kaupunki (2018) LinkkiData, Available at http://data.jyvaskyla.fi/data.php
Manchester: TfGM (2018) GM Public Transport Schedules - GTFS, Available at https://data.gov.uk/dataset/c96c4140-8b6c-4130-9642-49866498d268/gm-public-transport-schedules-gtfs
Paris: Île-de-France Mobilités (2018) Open Data, Available at https://opendata.stif.info/explore/dataset/offre-horaires-tc-gtfs-idf/
Stockholm & Uppsala: Trafiklab (2018) GTFS Sverige 2, Available at https://www.trafiklab.se/api/gtfs-sverige-2
Tampere: ITS Factory (2018) Tampere Public Transport GTFS feed, Available at http://data.itsfactory.fi

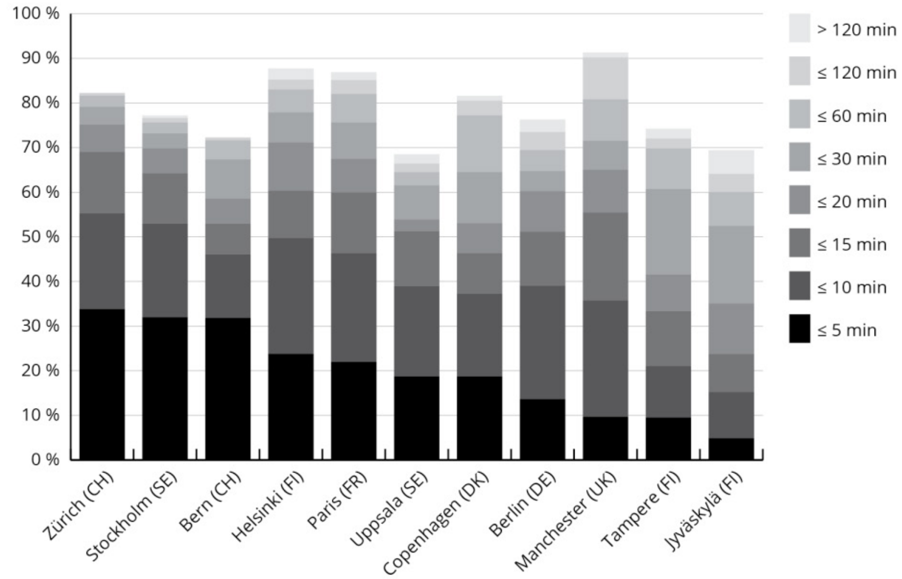

Fig. 4 Public transport reach in selected urban areas. The columns are sorted by the reach of headway of 5 minutes or less.

It should be noted that each of these benchmarked cities have a single, integrated ticketing system – except for Manchester, where each (market-based) operator has its own ticketing system. In Manchester, a third party provides integrated tickets, but these tickets are more expensive than single operator tickets (Palonen 2017). Therefore, to get access to full coverage in Manchester, the passengers need to pay a relatively higher price.

This method also allows analysing the level of service per operator. The total public transport reach and reach of the three biggest public transport operators in Manchester is shown in Fig. 5; individual operator services cover at best less than half of total inhabitants.

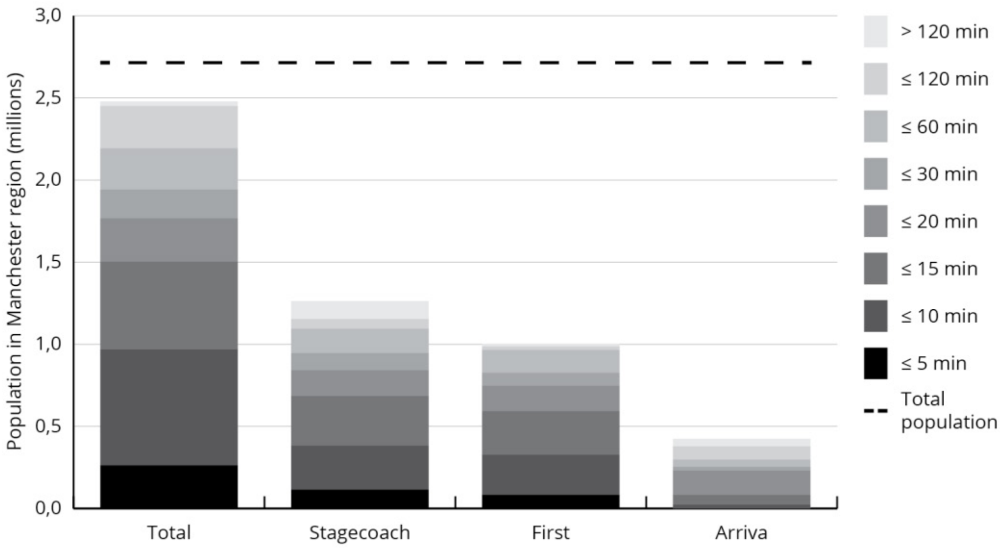

Fig. 5 Public transport reach per selected operators and in total in Manchester.

The problem with the reach of public transport in Manchester can be seen more clearly in Fig. 6. First Greater Manchester has good coverage and service in northern parts of the city while Stagecoach has equally good coverage and service in the southern parts of the city. In total, most of the city has good public transport service, but for passengers there are only few crosstown lines that do not require a transfer between operators.

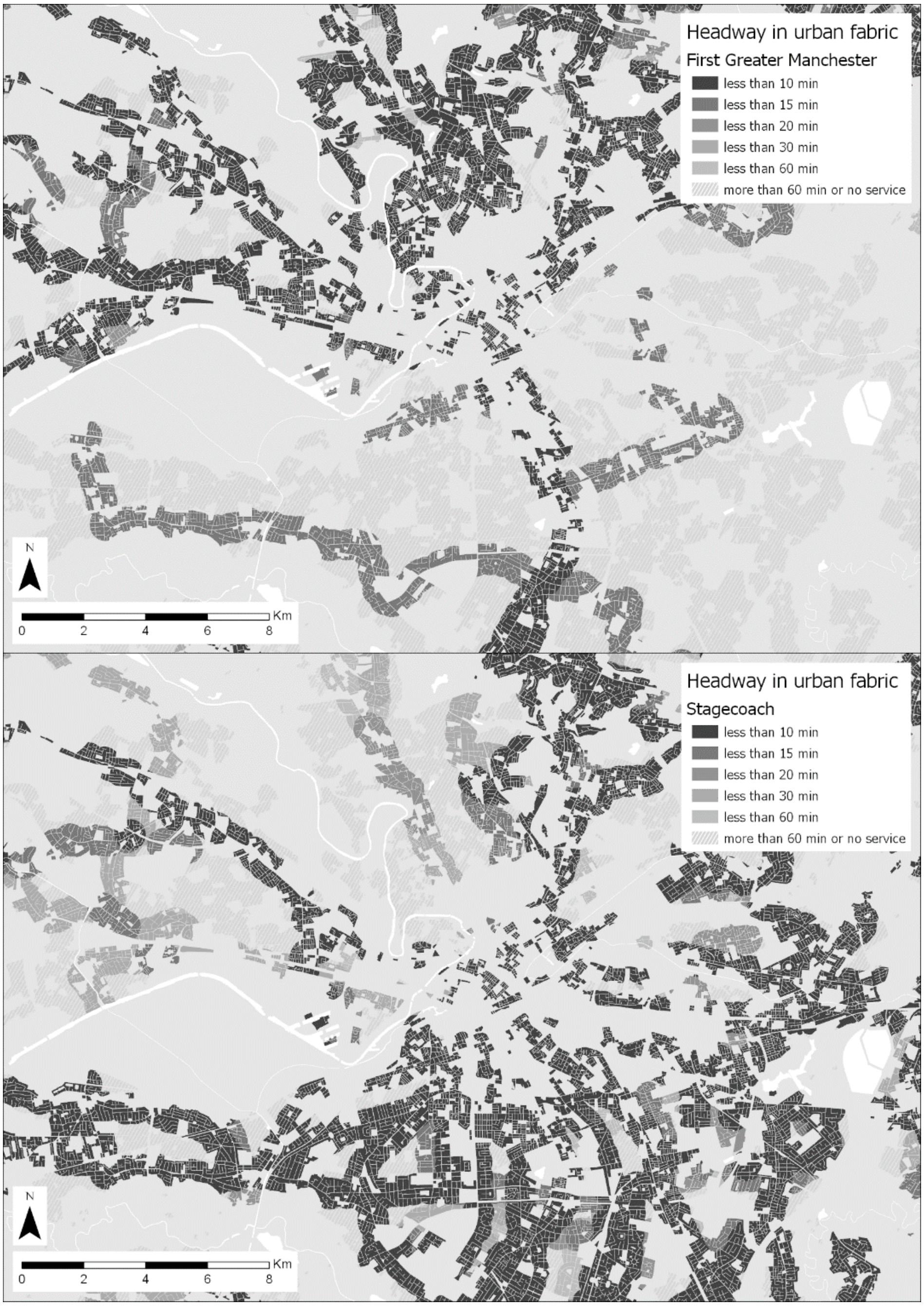

Fig. 6 First Greater Manchester (above) has good service coverage in the northern parts of Manchester, while Stagecoach Manchester (below) has good service coverage in southern parts of Manchester.

4. Discussion

All cities have not (yet) published their public transport timetable data as open data. However, establishing the Single European Transport Area includes the implementation of National Access Points (NAP) where EU countries should list the interfaces for accessing transport data, including public transport timetable data[6]. While the timetable data might not be in GTFS format in a NAP, it can still be used in this method, given that it is parsed and filtered appropriately. As the implementation of national access points continues, more cities can be benchmarked using the method presented here. This method can be also used for benchmarking cities outside Europe, given that appropriate population grid and land use data is available.

Here all the public transport stops had the same size catchment area. Transit Cooperative Research Program (2003) suggests that passengers are likely to walk longer distances to rail transport, i.e. larger catchment areas could be assigned to rail transport stops. This method allows the altering of catchment areas by transport mode, it could be utilized in further studies.

Besides core timetable data, GTFS schema includes also information on whether routes and stops are accessible by wheelchair. Using this data, it is possible to provide public transport level of service benchmark considering (wheelchair) accessibility.

This method takes into consideration only the population reach of public transport. If workplace density data were available (as open data), it could provide supplementary information on the public transport level of service considering the reachability of workplaces. Further development could also include network connectivity study and the reachability of public services and certain commercial services by public transport. Going further, the method could be utilized e.g. in transport accessibility and transport poverty studies.

Acknowledgement

Major part of this work was done while both authors were affiliated with Transport Research Centre Verne at Tampere University of Technology (currently Tampere University).

References

EFGS (European Forum for Geostatistics, 2011) GEOSTAT 1A – Representing Census data in a European population grid.

Finnish Transport Agency (2015) Joukkoliikenteen palvelutason määrittely (in Finnish; Defining the Public Transport Level of Service), Helsinki.

Google (2018a) GTFS Static Overview. Google Developers. Available at: https://developers.google.com/transit/gtfs/ Cited: 21.8.2018

Google (2018b) General Transit Feed Specification Reference. Google Developers. Available at: https://developers.google.com/transit/gtfs/reference/ Cited: 21.8.2018

Mulley, C., Ho., C., Ho, L., Hensher, D., and Rose, J. (2018) Will bus travellers walk further for a more frequent service? An international study using a stated preference approach, Transport Policy, 69 (2018) 88–97

OECD (2013). Definition of Functional Urban Areas (FUA) for the OECD metropolitan database.

Palonen, T. (2017) Tällaista tuli tällä rahalla – Kaupunkiliikenteen järjestämistavan vaikutus palvelutasoon (in Finnish; You Get What You Pay for – Level of Service in Publicly Procured and Market Based Urban Public Transport), Tampere University of Technology (Master's Thesis), Tampere.

Poelman, H., Dijkstra, L. (2015) Measuring access to public transport in European cities, European Commission Working Papers.

TCRP (Transit Cooperative Research Program, 2003) Transit Capacity and Quality of Service Manual, 2nd Edition, Transportation Research Board, Washington, D.C.

Walker, J. (2011) Human Transit, Island Press, Washington, D.C.

[6] Commission Delegated Regulation (EU) 2017/1926 of 31 May 2017 supplementing Directive 2010/40/EU of the European Parliament and of the Council with regard to the provision of EU-wide multimodal travel information services