Public transport accessibility for Great Britain: An open dataset

Rafael Verduzco, MSc and Dr. David McArthur
June - 2022
Overview

- Public transport accessibility indicators are useful to understand a variety of spatial phenomena such as (Levinson, 2021):
  - Risks of unemployment,
  - Land prices,
  - Travel behaviour,
  - Gentrification, among many others...

- Comprehensive accessibility measures demand considerable resources, i.e. computational, time, and technical

- The above can represent barriers for some applied researchers

- The present project aims to bridge this gap

- This is the first attempt which covers the whole GB using a single model
Open Science and reproducibility

- Although there is good support for the idea, reproducibility levels are not improving in the field of GIS (Nüst et al., 2018)

- Our motivation:
  - Transparency
  - Increased potential for (collective) improvement
  - Possible collaborations

- Our challenge: Make it as open and reproducible as possible
  - All software open-source
  - All inputs open access
  - Outputs freely available
  - Limitations: distributing all components (e.g. data inputs)
  - Still, the details and sources are provided in a technical report
Accessibility

The extent to which land-use and transport systems enable (groups of) individuals to reach activities or destinations by means of a (combination of) transport mode(s)

Geurs & van Wee (2004)
Accessibility

There exist a number of accessibility classes, i.e.: (Geurs & van Wee 2004):

- Infrastructure-based (focuses on mobility);
- Location-based (includes mobility and land use);
- Person-based (focus at the individual level), and;
- Utility-based (from economics perspective).

Location-based measures are the most frequently used given their:

- Consistent theoretical specification
- Flexibility
- Ease of interpretation (in some cases)
Location-based measures

- Generic specification:\(^1\)

\[ A_{ik} = \sum g(W_{jk}) f(c_{ijk}) \]

- Example, cumulative opportunities:
  - Number of jobs that can be reached in 30 minutes using public transport.

- For this work:
  1. Cumulative opportunities various time cuts, i.e. from 15 to 120 every 15 minutes
  2. Relative cumulative opportunities (15 to 120 minutes)
  3. Dual/minimum travel time

Note: 1, Páez et al., 2012; Levinson & Wu, 2020.

www.ubdc.ac.uk
## Origins

- 2011 LSOA/DZ’s population weighted centroid (ONS, Scottish Government)

<table>
<thead>
<tr>
<th></th>
<th>England</th>
<th>Scotland</th>
<th>Wales</th>
<th>GB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Count (N)</td>
<td>32 844</td>
<td>6 976</td>
<td>1 909</td>
<td>41 729</td>
</tr>
<tr>
<td>Population 2020 (mean)</td>
<td>1 722</td>
<td>784</td>
<td>1 660</td>
<td>1 562</td>
</tr>
<tr>
<td>Surface area sq. km. (Mean)</td>
<td>4.0</td>
<td>11.2</td>
<td>11.1</td>
<td>5.6</td>
</tr>
</tbody>
</table>
## Destinations

<table>
<thead>
<tr>
<th>Destination</th>
<th>England</th>
<th>Scotland</th>
<th>Wales</th>
<th>GB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Employment (millions)</td>
<td>26.3</td>
<td>2.5</td>
<td>1.3</td>
<td>30.0</td>
</tr>
<tr>
<td>GPs</td>
<td>6,560</td>
<td>922</td>
<td>405</td>
<td>7,887</td>
</tr>
<tr>
<td>Hospitals</td>
<td>1,174</td>
<td>246</td>
<td>90</td>
<td>1,510</td>
</tr>
<tr>
<td>Education: Primary schools</td>
<td>16,608</td>
<td>2,003</td>
<td>1,242</td>
<td>19,853</td>
</tr>
<tr>
<td>Education: Secondary schools</td>
<td>2,893</td>
<td>359</td>
<td>205</td>
<td>3,457</td>
</tr>
<tr>
<td>Urban centre: Subcentre</td>
<td>336</td>
<td>50</td>
<td>35</td>
<td>421</td>
</tr>
<tr>
<td>Urban centre: Main</td>
<td>146</td>
<td>23</td>
<td>13</td>
<td>182</td>
</tr>
<tr>
<td>Supermarkets</td>
<td>5,467</td>
<td>672</td>
<td>339</td>
<td>6,478</td>
</tr>
</tbody>
</table>
Travel time estimates – Software and sources

- Core component of accessibility measures. Also, the *bottleneck* for a long time.

- Routing software: R5 (Conway et al., 2017; Conway et al. 2018):
  - Implemented in R via the {R5R} package (Saraiva et al., 2021)
  - R5: Rapid Realistic Routing on Real-world and Reimagined networks
  - Open-source software
  - Explodes advantages of parallel of computing

- Main inputs (open data):
  - Road and pedestrian network: OpenStreetMap (OSM)
  - PT timetables:
    - Bus Open Data Service (BODS) for local services, and
    - Rail Delivery Group (ATOC) for heavy rail.
Travel time estimates – Model parameters

- Follows Journey Time Statistics (DfT, n.d.) where possible
- All-to-all travel time matrix, i.e. from each LSOA to all other.
- Approx. 95 million OD routes computed in ~84 hrs or 3.5 days

- Mode: combines *walking* and *public transport*
- Departure: Tuesday 22 of November 2021 at 07:00 a.m.
- Time-window: 3 hours
- Maximum journey: 120 minutes
- 3 rides maximum (0 min)
- Walking distance to access/egress unlimited
Travel time estimates – TT uncertainty

- Considering variability of travel time using percentiles (0-100), i.e.:
  - Low percentile – High flexibility of traveller (e.g. regional commuter)
  - High percentile – Low flexibility (e.g. turn-up and go)
# Outputs

- Open code (GitHub repository)
- Technical report (Zenodo repo)
- All-to-all travel time matrix (25, 50, 75 TT percentiles) (UBDC repo)
- Accessibility indicators for key services/amenities: (UBDC repo)

<table>
<thead>
<tr>
<th>Service</th>
<th>Cumulative</th>
<th>Relative</th>
<th>Minimum TT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Employment</td>
<td>**</td>
<td>**</td>
<td></td>
</tr>
<tr>
<td>GPs</td>
<td>**</td>
<td>**</td>
<td>**</td>
</tr>
<tr>
<td>Hospitals</td>
<td>**</td>
<td>**</td>
<td>**</td>
</tr>
<tr>
<td>Education: Primary schools</td>
<td>**</td>
<td>**</td>
<td>**</td>
</tr>
<tr>
<td>Education: Secondary schools</td>
<td>**</td>
<td>**</td>
<td>**</td>
</tr>
<tr>
<td>Urban centre: Subcentre</td>
<td></td>
<td></td>
<td>**</td>
</tr>
<tr>
<td>Urban centre: Main</td>
<td></td>
<td></td>
<td>**</td>
</tr>
<tr>
<td>Supermarkets</td>
<td>**</td>
<td>**</td>
<td>**</td>
</tr>
</tbody>
</table>
Accessibility to employment in GB

90 minutes

120 minutes

Greater London

Employment (%)

- 0.00-0.52
- 0.52-1.22
- 1.22-2.02
- 2.02-2.84
- 2.84-3.77
- 3.77-4.91
- 4.91-6.19
- 6.19-7.69
- 7.69-9.56
- 9.56-11.49
- 11.49-13.27
- 13.27-14.94
- 14.94-16.65
- 16.65-18.42
- 18.42-20.23
- 20.23-22.13
- 22.13-24.28
- 24.28-26.73
- 26.73-29.56
- 29.56-34.52

Cities:
- London
- Manchester
- Glasgow
- Newcastle
- Birmingham
Accessibility to employment in Scotland

Central Belt, Scotland

Glasgow

Edinburgh

Aberdeen

Dundee
Accessibility to education: Primary schools

15 minutes

West (midlands) England

Birmingham

Manchester

Liverpool

London

Manchester

Glasgow

Newcastle

Birmingham

London

Nottingham
Limitations

• Transport model: Intercity coaches are not included (format not compatible)
• Location of destinations are represented by LSOA/DZ centroids
• Heterogeneity of geographic areas, i.e. LSOA vs DZ
• Heterogeneity of input sources, e.g. schools, hospitals.
• Basic measures, e.g. they can account for demand or other details.
Future work

- Other modes:
  - Walking or cycling. Possible to account for topography or sensitivity of travellers to type of road, for example.
  - Driving, useful to set a benchmark.
- Alternative scenarios, e.g. Covid vs post-Covid.
- Higher spatial resolution, e.g.:
  - Origins as output area (OA), or
  - Destinations at the point level.
- Visualization of data, e.g. regional maps and statistics, comparison between major cities, or isochrones.
- Interactive dashboard
Final thoughts

- Provide a contribution to researchers to various degrees/levels:

1. Directly applicable using readily available indicators
2. Customizable measures, e.g. using TT and code for other type of measures, consider demand, etc.
3. Specialized researchers can draw on basic code to develop alternative scenarios, measures of uncertainty, or use it as benchmark, etc.
Thank you!

Contact

Email: j.verduzco-torres.1@research.gla.ac.uk
Twitter: @raffverduzco

Links to resources

- Open code: https://github.com/rafavidz/access_uk
- Technical report: https://zenodo.org/record/6759240#.Yrwg0nbMKUl
References