ITF Roundtable – Big Data and Transport Models

Use of mobile telecommunication data in transport modelling – a French case study

15th December, 2020

Marie ARBOUET – Expert in transport economics (SYSTRA)
Thibault JANIK – Consultant in transport economics (SYSTRA)
Context

- Current models mainly based on household surveys
- Known limits of household surveys
- Opportunity to work with additional data: mobile phones data
- Widespread usage of mobile phones in the population
- Technologies available for tracking
- Privacy rules
Objectives and Datasets

Objectives set by SNCF Réseau:

- Explore the potential of mobile phone data
- Propose sensitivity tests
- Validate the results against independent data sources

Datasets provided and adjusted by the operator:

- Trip volumes observed for each day of November and December 2018
- Annual average day trip volumes for 2018
- Annual average of trip volumes during basic working days
Zone systems

Fine zone and Coarse zone systems

Source: Systra (2019)
Immobility threshold

Influence of the immobility threshold on trip numbers

Coarse zone system

Fine zone system

Source: Systra (2019)
## Quality indicators

<table>
<thead>
<tr>
<th>Indicators</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>QI1</strong></td>
<td><strong>Influence of statistical adjustment</strong>&lt;br&gt;This indicator indicates whether the flow has undergone a significant statistical adjustment to be representative of the population or not.</td>
</tr>
<tr>
<td><strong>QI2</strong></td>
<td><strong>Modal assignment</strong>&lt;br&gt;This indicator is high when a large part of the traces can be confidently attributed to a mode.</td>
</tr>
<tr>
<td><strong>QI3</strong></td>
<td><strong>Precision of the origin and destination areas</strong>&lt;br&gt;This indicator relates to the frequency of mobile locations. When this frequency is high, the reliability of locating the origin and destination area is better.</td>
</tr>
<tr>
<td><strong>QI4</strong></td>
<td><strong>Terrestrial modal assignment</strong>&lt;br&gt;This QI focuses only on terrestrial modes. The terrestrial modal assignment is largely determined by the passage through cells which are characteristic of a given mode.</td>
</tr>
</tbody>
</table>
Data validation

Comparison of Paris-Province and Province-Province flow between GSM traces and the model

Source: Systra (2019)
Conclusion

- Complexity of the data
- Set a zoning system and a immobility threshold adapted to the scope of analysis
- Favour the datasets over a long period
- Transport mode reliability
- An incredible source of information for the transport analysis and modelling
Cell-towers coverage

Cell-towers coverage – all operators – France

Source: Systra (2019)
Quality indicator

Distribution of QI 2 values by distance class

Source: Systra (2019)
GSM traces vs Air data

Comparison of the main radial air flows 2018 GSM / DGAC

Source: Systra (2019)
GSM traces vs Eurostat data

Comparative of rail flows Region * Region – GSM traces / Eurostat 2015

Source: Systra (2019)