United Nations ECE/TRANS/WP.6/2021/11

Economic and Social Council

Distr.: General
25 March 2021

Original: English

Economic Commission for Europe

Inland Transport Committee

Working Party on Transport Statistics

Seventy-second session
Geneva, 9–11 June 2021
Item 2 of the provisional agenda
Big Data for transport statistics

Using floating car data to generate transport statistics

Note by the secretariat

Summary

This document describes the experience of Germany in utilising floating car data, a potential new data source to use for the generation of passenger journeys. Results indicate that origin-destination data generated from the floating car data are comparable to those collected by a traditional mobility survey. The annex gives more technical details of the methodology of the study.

I. Background

1. As explored in recent webinars co-organised by the secretariat and the International Transport Forum (ITF) (ECE/TRANS/WP.6/2021/4), the COVID-19 epidemic has accelerated the transition from survey-based transport statistics towards utilising new data sources, including big data sources. One such source is floating car data (FCD), which refers to data collected directly by the vehicle itself as it is in motion, typically covering its location and speed. The remainder of this document sets out the experience of Germany in this area.

II. Summary of the experience of Germany on floating car data

2. The utilization of FCD, which are generated in large quantities and almost continuously during logistics and navigational processes, poses a new challenge to traffic research. The vast amount of data creates the expectation that these data will make it possible to provide information for traffic planning quicker and more extensively in the future, and that this could reduce the costs of conducting conventional empirical surveys. The German Federal Ministry of Transport and Digital Infrastructure (BMVI) therefore commissioned the Bergische Universität Wuppertal to carry out the research project "Nationwide OD-Matrices 2015 for Car and Truck Transport – Analysis based on satellite data" (FE 97.372 / 2016).

3. FCD, which are generated with navigation devices and are based on spatially very precise satellite positioning systems, make it possible to obtain information about the start and destination locations as well as speeds travelled and routes of vehicles. This information can be analyzed in different temporal and spatial contexts. The goal of the research project was to clarify whether and, if so how, valid nationwide origin-destination (OD)-matrices can be estimated for an entire year.

4. The investigation was based on an FCD data set for 2015, which included data from several navigation service providers containing both passenger and freight vehicles. This data set was examined for spatial and temporal clusters in multi-level evaluations. The vehicle type assignment was checked and improved and, on this basis, the reported spatial positions (signal chain) were segmented into trips. Then the trip purposes were assigned by means of existing empirical studies through similarity analysis, but only for passenger car data; a comparable assignment of trip purposes was not possible for freight transport.

5. Since the trips reported with FCD only comprised a small sample of all nationwide trips in a full year and as a selective use of navigation devices was recognizable in the data set, the OD-matrices for car transport were extrapolated using the parameters of the Mobility in Germany survey (Mobilität in Deutschland (MiD)). For this purpose, only the marginal sums of the OD-matrix were used, so that the distribution of the total traffic to the individual relations was mapped exclusively on the basis of the FCD database. As no suitable methods for estimating the marginal sums were available for the goods data, the OD-matrix of truck transport was extrapolated with the help of the traffic volumes at the permanent counting stations of the federal highway road network.

6. The subsequent analyses of the extrapolated matrices were based on aggregated parameters (mileage, number of journeys, distance travelled) and showed overall satisfactory results for the car matrix estimated with FCD. It should be noted that all comparative values used were also generated from random samples and model calculations. In both matrices, cross-border and transit trips could not be mapped representatively due to the strong spatial clusters of the FCD. For truck transport, the extrapolation suffers from the lack of validation of the marginal sums. Short journeys without passing a permanent counting station made it difficult to estimate valid extrapolation factors for the relevant relations.

7. Despite the identified problems, the FCD data is considered to be a promising source of information for large-scale origin-destination links, which would be hard to obtain from a survey. A further increase in the use of navigation devices is to be expected in the future, especially with regard to shorter and regular journeys. In addition, the harmonization of the European mobile communications market will remove barriers which, at the time of the investigation, still prevented the use of smartphone-based navigation services and thus contributed to the selective mapping of cross-border journeys. There is also a considerable need for further methodological research into the utilization of FCD.

III. Conclusions

8. The results of the German experience with FCD shows that the data source may have potential to either complement or even replace traditional mobility survey data in the future. In addition, improvements in tracking cross-border journeys with FCD are to be expected in the future due to changes in European mobile communications market. Further work is required on ensuring that FCD data can accurately cover the entire population.

Annex

Methodology details of the project

The objective of the project was to use an FCD data set from 2017 to verify and validate the method (namely interdependency matrix estimation) developed in the preceding research project FE 97.372/2016 "Nationwide Traffic Interdependency 2015 in Motorized Individual Transport (MIV) – Analyses Based on Satellite-Based Data".[1] Thus the research question was to what extent the distribution of traffic remains stable over time and whether the data can be used for the estimation of nationwide traffic interdependencies, that is origin and destination patterns. Within the framework of the research project, the developed methods were applied to a floating car data set of the German Automobile Club (Allgemeiner Deutscher Automobil-Club e.V. (ADAC)) from the year 2017. Already existing algorithms for trip segmentation and extrapolation were further improved during the project.

The underlying data volume increased by about 30 per cent compared to the year 2015. Due to the hybrid extrapolation procedure, which is based on the region type, distance class and purpose of the journey dependent indicators generated by the Mobility in Germany survey (Mobilität in Deutschland (MiD)) 2017, the generated indicators e compare favourably. The evaluation has shown that neither the spatial nor the temporal distribution of the data has changed fundamentally. The additional journeys depicted by the increased data volume are distributed analogously. The destination selection behaviour as well as the symmetry of the generated FCD trips correlate very strongly between the individual years for both motorized individual traffic (MiV) and heavy traffic (SV). The study also showed that despite the abolition of Europe-wide mobile roaming charges, the mapping of cross-border journeys is not yet sufficient to obtain fully conclusive results. In addition to the cell divisions at NUTS 3[2] and LAU 1[3] level used for the matrix generation, an INSPIRE-compliant[4] dynamic grid cell model based on the 100 x 100 m grid cells of the 2011 census was developed. Based on previously defined parameters for source and destination traffic, the method allows the dynamic generation of FCD-based grid cells that represent a multiple of the census grid and thus allow a much more data-affined cell division.

---

[1] Available in German: https://www.bmvi.de/SharedDocs/DE/Anlage/G/verkehr-in-zahlen-praesentation-3.pdf?__blob=publicationFile.

[2] Referring to Nomenclature of Territorial Units for Statistics of the European Commission. NUTS 3 level is approximately analogous to municipal level; there are 429 NUTS 3 regions in Germany.

[3] Local Administrative Units, again defined by the European Commission. LAU 1 is the level below NUTS 3.

[4] INSPIRE is the European Union spatial data infrastructure.