

Estimating the Residential Population from Mobile Phone Data, an Initial Exploration*

Benjamin Sakarovitch, Marie-Pierre de Bellefon, Pauline Givord and Maarten Vanhoof

Key Question

Mobile phone data provide recordings with high spatial resolution and temporal frequency. They are considered a promising source for measuring people's mobility and visitor numbers at very fine spatial or temporal levels. Their use for statistical purposes raises a number of issues (quality of the information collected, representativeness of the available data, necessary adjustments), that are examined with a focus on the estimation of indicators of resident populations.

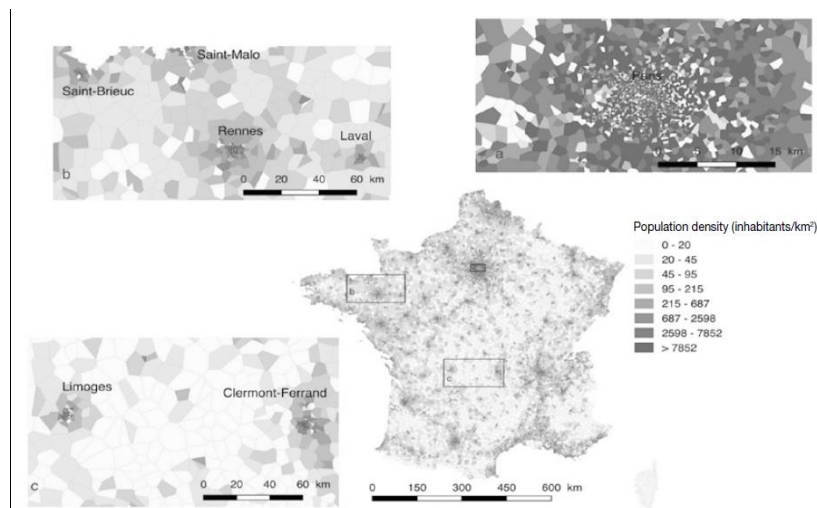
Methodology

The place of residence of each individual user is assigned, on the basis of anonymised data provided by their mobile phone operator over five months in 2007, according to the area covered by the most recurring antenna used for call transmission. The resident population constructed in this way is then compared to that provided by geolocalized tax data.

Main Results

- The resident population estimates based on mobile telephony differ significantly from the reference population provided by the tax source.
- The modelling of antenna coverage (Voronoi tessellation) and then the transposition of resident population estimates to the communal grid explain most of the differences observed.
- The estimates are more precise at the communal level than at the level of the antenna coverage area - and almost as much as at a larger mesh (employment area, department).
- Adjusting for the local heterogeneity of the operator's market shares improves the quality of the estimate, but requires auxiliary information such as a customer file, the availability and reliability of which are not guaranteed.
- The gaps between the estimated and reference resident population have a pronounced spatial structure.
- However, mobile phone data seem promising for the study of seasonal fluctuations in population.

Population Density per Municipality Calculated from Mobile Phone Data



Note: Estimates adjusted at the departmental level using the client file.

Sources: CDR, customer file and Filosofi; authors' calculations.

Message

The use of mobile phone data for population counts raises several methodological issues. The modelling of the antenna coverage area, the choice of the analytical grid and the location of events on this grid play a decisive role, as well as the methods used to extend the estimates based on an operator's customers to the general population. Using data from recordings of subscribers' location at higher frequency than billing data could provide greater precision for the detection of anchor points (such as the place of residence). The study of the presence of population and spatial and temporal variations in population density are among the most promising applications for official statistics.