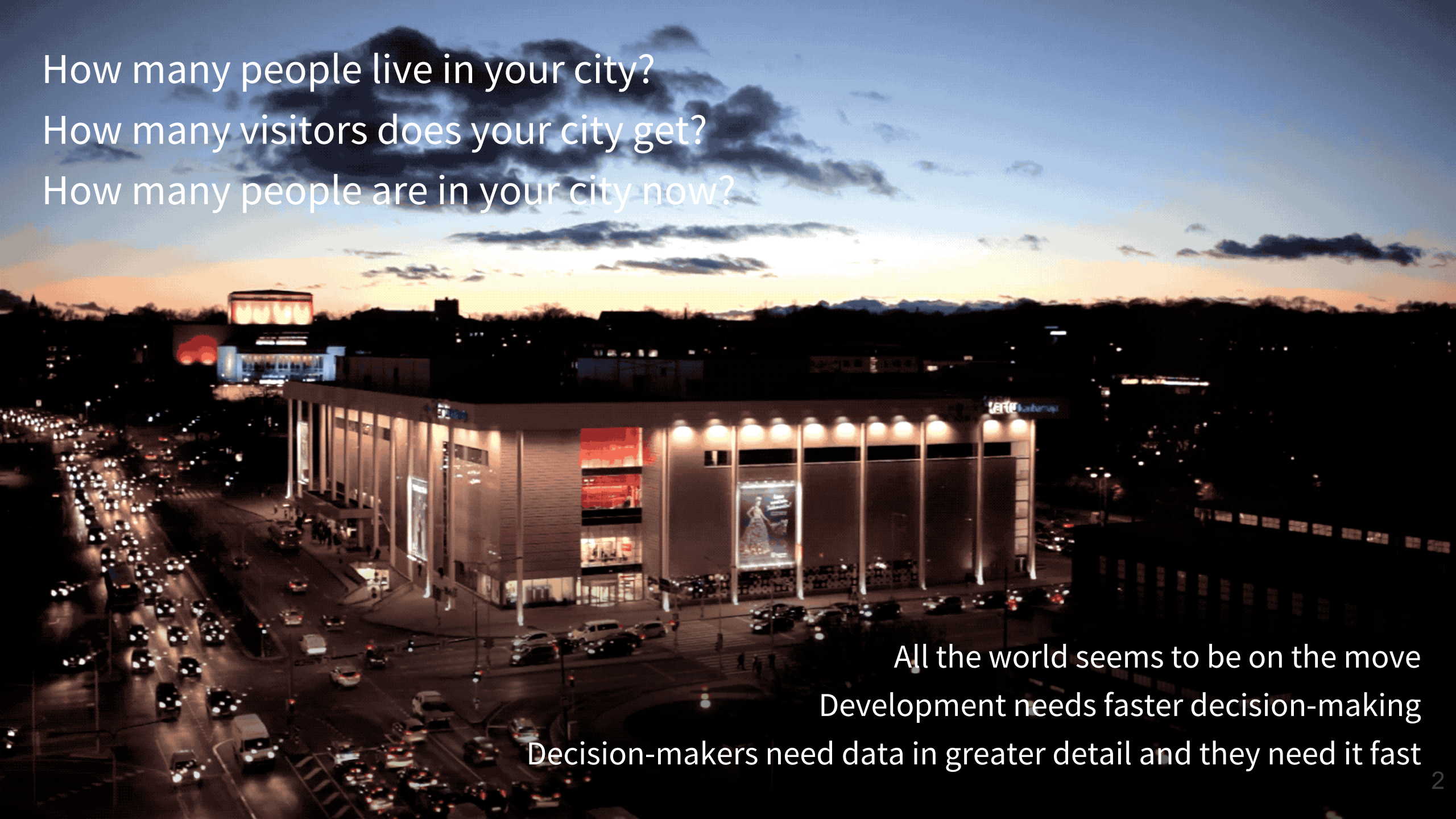
Guide on Mobile Phone Data (MPD) for Dynamic Population Mapping

UN Committee of Experts on Big Data and Data Science for Official Statistics has prepared guiding materials for the use of mobile positioning data for dynamic population mapping for the statistical community

Presenter: Siim Esko, Positium, Estonia
siim.esko@positium.com

Author/s: Siim Esko (Positium); Tracey Li (Flowminder Foundation); Ayumi Arai (University of Tokyo); Margus Tiru (Positium); Kaisa Vent (Positium); Titi Kanti Lestari (BPS Statistics Indonesia); Sabrina Juran (United Nations Population Fund); Dunstan Matekenya (World Bank)


How many people live in your city?
How many visitors does your city get?
How many people are in your city now?
All the world seems to be on the move
Development needs faster decision-making
Decision-makers need data in greater detail and they need it fast

Model Reality in Statistics

Objective: Make a data model of the real mobility of people using spatial data

Challenges of Current Population Methods

Census

- **Cost** pressures
- **Timeliness** and accuracy concerns
- Response **burden**: intrusiveness and reduced cooperation issues
- **Inclusivity** concerns: unsafe areas, difficult living arrangements (nomads, homelessness, closed communities)

Registries

- **Needs a good quality system** of statistical registers
- **Timeliness** of administrative registers
- Difficulty in identifying sub-populations or obtaining **small area data**
- **Under-coverage** errors - e.g. illegal or unregistered migrants are missing
- **Over-coverage** errors - e.g people haven’t officially declared their emigration

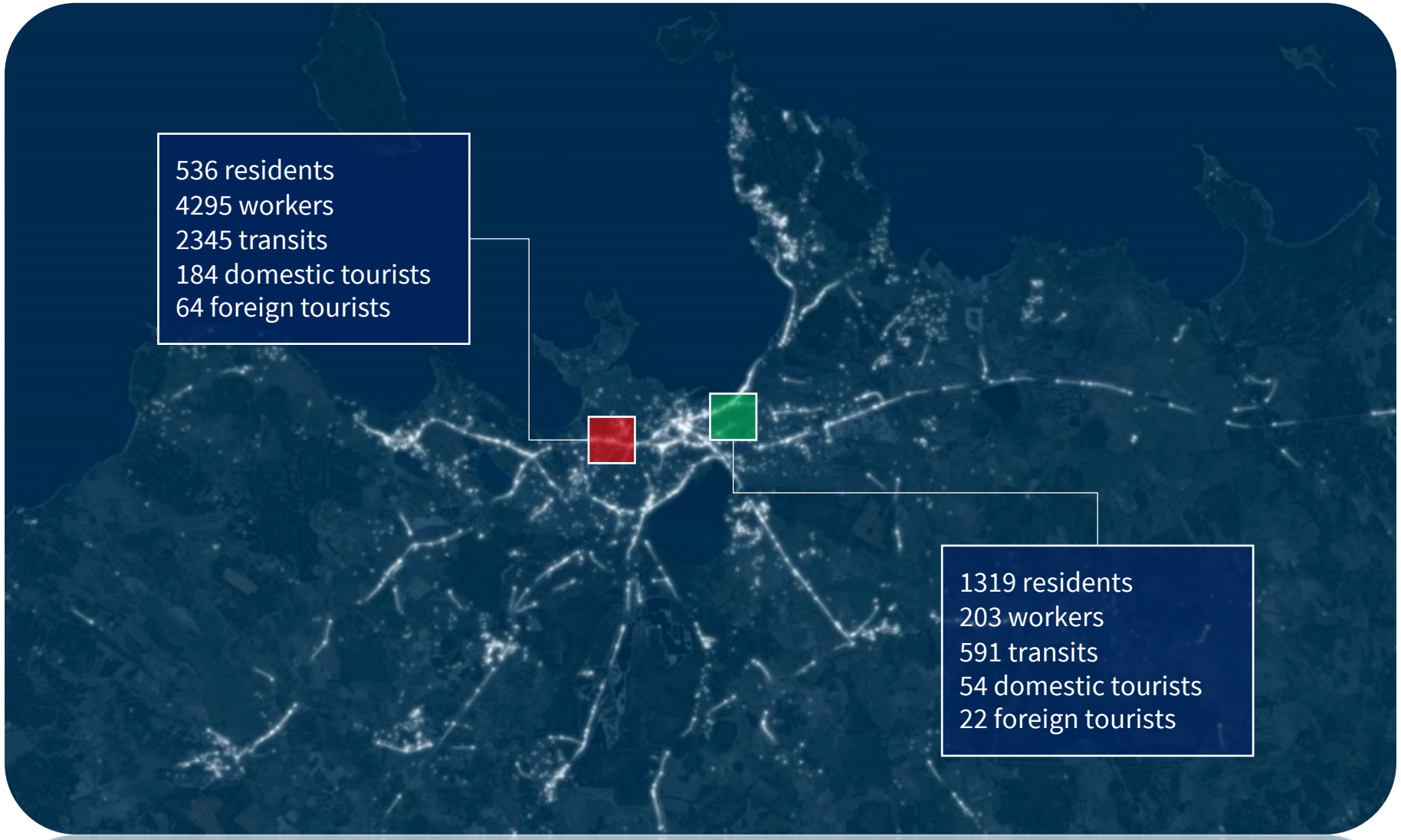
Dynamic Population Mapping aims to answer

how to **map population dynamically** - without being dependent on logistics of surveys or the census

&

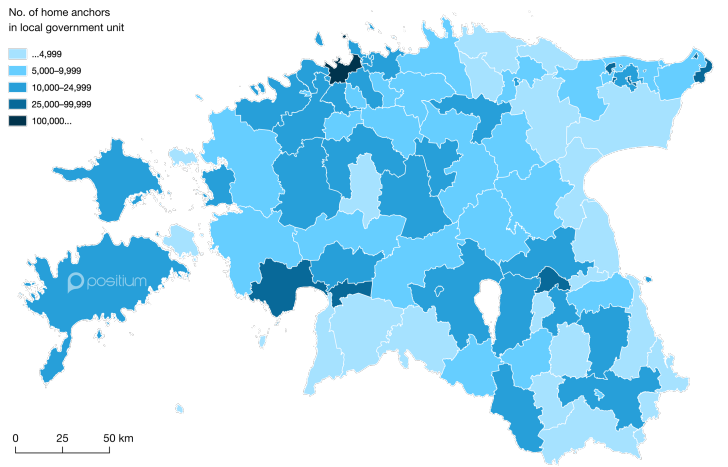
how to **map a dynamically-behaving population** - understanding de facto presence at any time, even away from place of residence

Dynamic Population Mapping Based on MPD

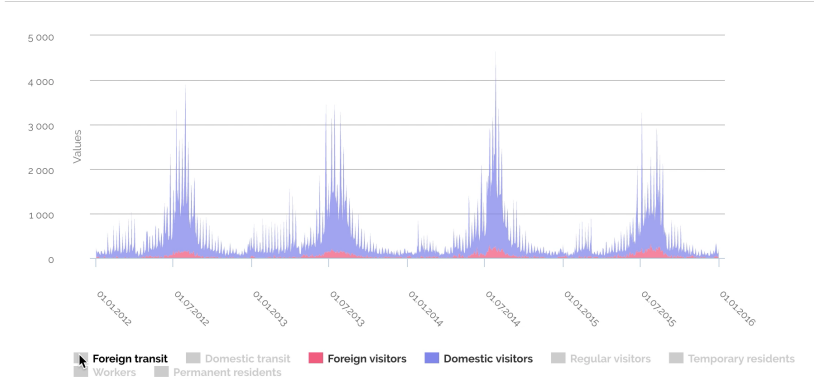


Definitions Used

***de jure* population** mapping measures the **presence of places of residence** of subscribers in the geographical area



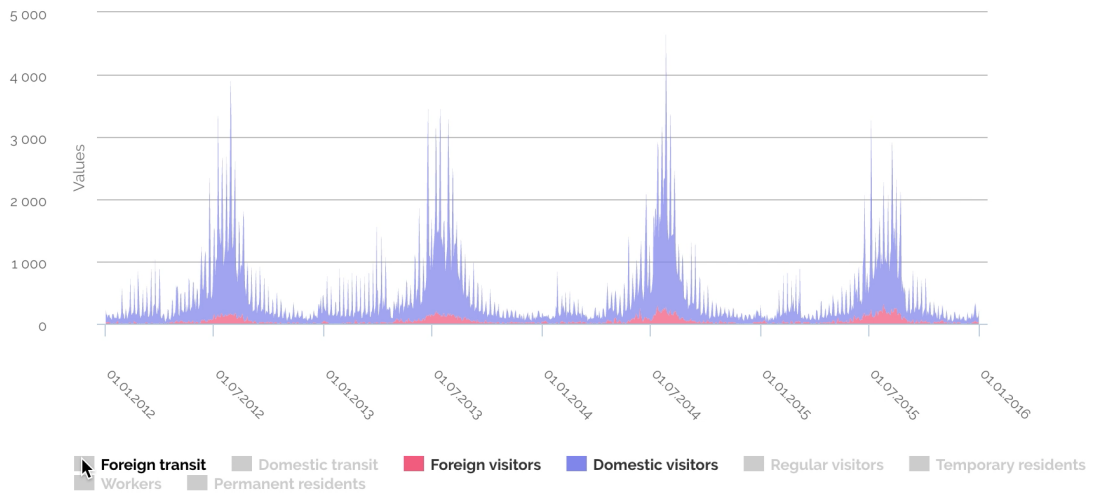
***de facto* (dynamic) population** mapping measures the **presence of subscribers** in the geographical area



8+ Uses of MPD for Population

Guide on the Use of Mobile Phone Data for DYNAMIC POPULATION MAPPING

United Nations | DESA Statistics Division



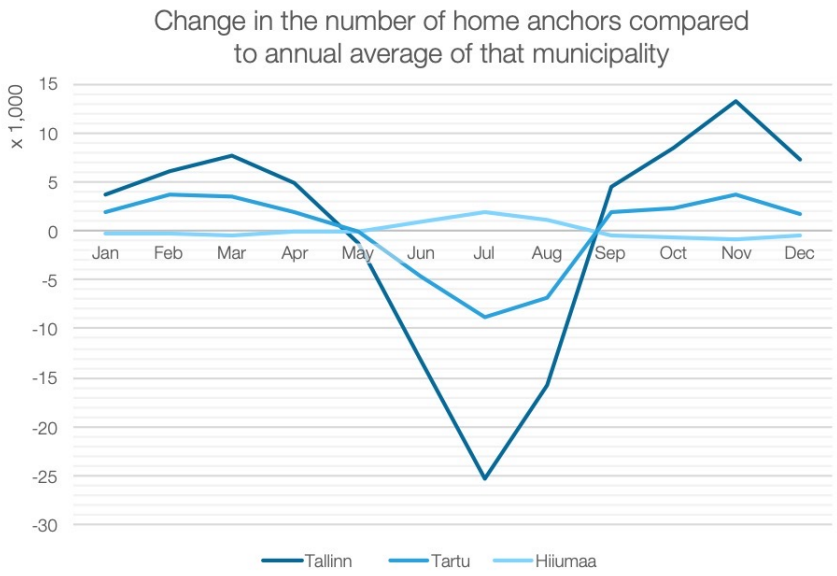
1. Resident population mapping
2. Daytime population mapping
3. De facto population mapping
4. Monitoring population redistributions caused by COVID-19 mobility restrictions
5. Infrastructure and resource planning
6. Creating dynamic sample frames for surveys
7. Census
8. Disaster preparedness planning and response

...

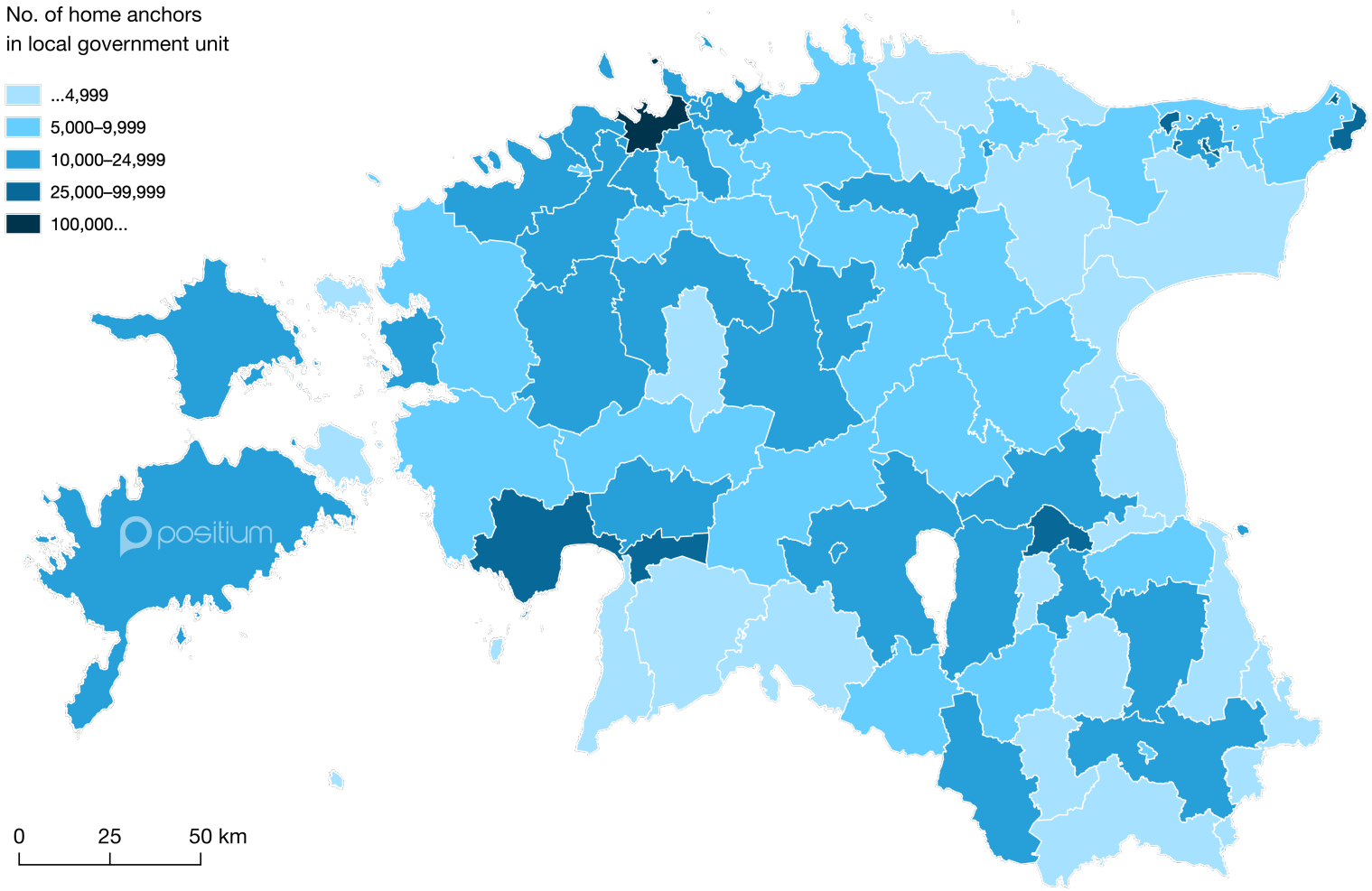
Resident population

Number of residents in local government units

When number of residents falls in Tallinn and Tartu over summer, it increases on the islands, such as Hiiumaa. In Tartu, there are many university students, who leave during summer.



Annual average number of home anchors: Tallinn 423,204; Tartu 96,133; Hiiumaa 11,320



Daytime Population

Mobile positioning data

▼

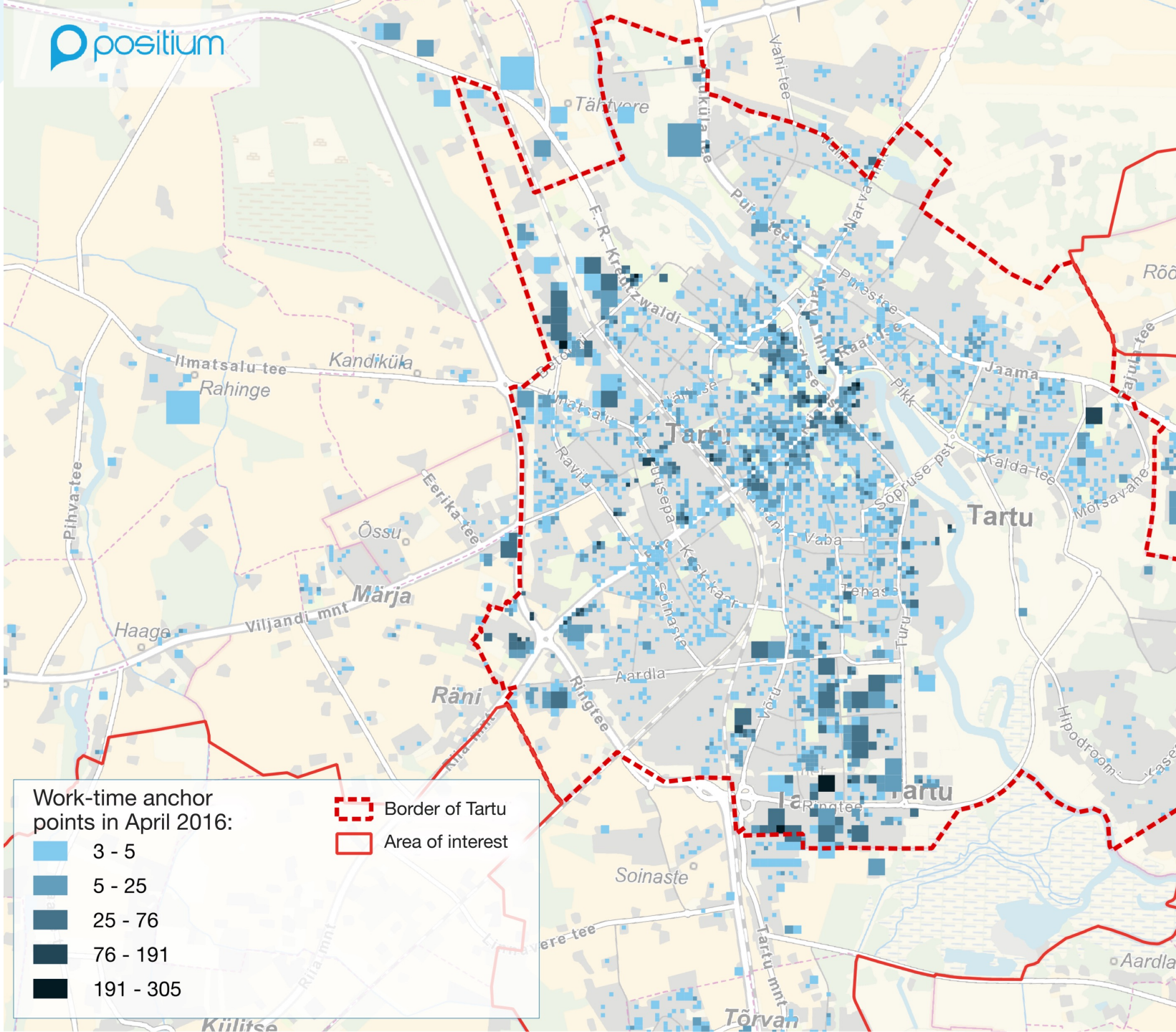
Cleaned and processed

▼

Work-time anchors detected

▼

Aggregated to adaptive grid



Census

Opportunities exist to complement the census during:

- **Census preparation and execution** (establishing survey frames)
- **Intercensal period** (population projections)
- **Measuring possible undercounted populations** (nomadic populations)
- **Adding features** (daytime population, migration or commuting)

And to carry out the Census through innovative combination of technology:

- **Data driven census** (combining administrative and big data)

Methodological challenges

And some methods to overcome them

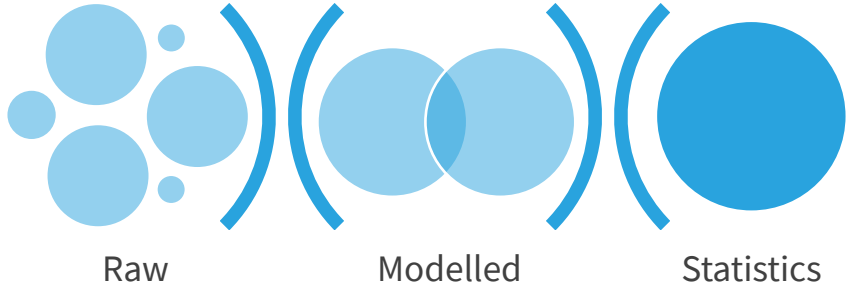
Main Methodological Questions

- How to build the right data model?
- How to detect place of residence and select the right home detection algorithms?
- How to ensure data coverage and representativity?
- How to model population density?
- How to validate the results?

Choice of model

Simplified model

- Statistical calculation done directly on raw data
  - E.g. simple SIM card counts or tower density measurements in a grid
- Quick indicators
- Modelling done post-processing
- Not suitable for official statistics



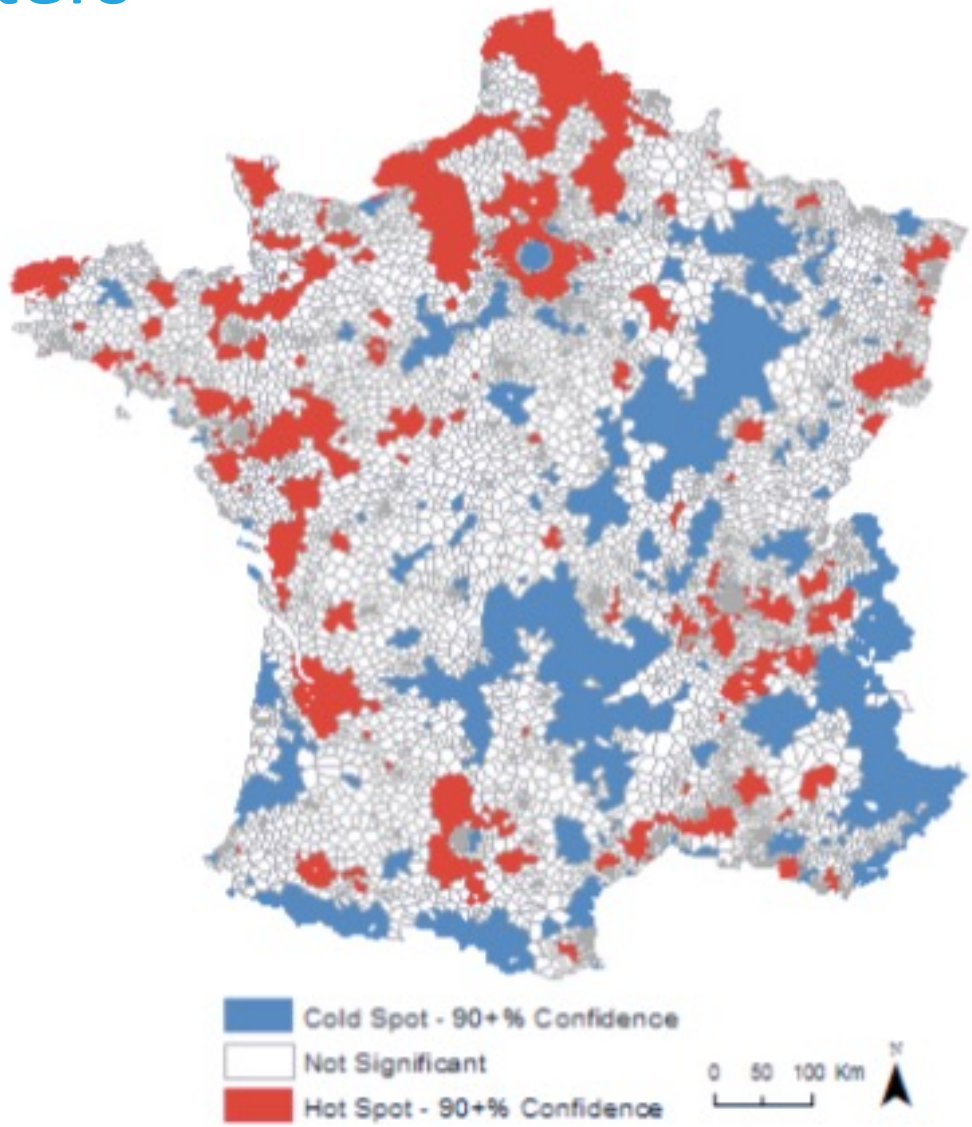
General model

- A model of reality is built for each subscriber
- Data model matches official definitions
- Statistical concepts are applied in late data processing
- Allows combining and comparing results for many domains
- Most useful for official statistics

Place of Residence Algorithm Matters

Choice of criteria in home detection algorithms influences the results significantly

- up to 40% in France.



Vanhoof, M et al (2018) Assessing the quality of home detection from mobile phone data for official statistics

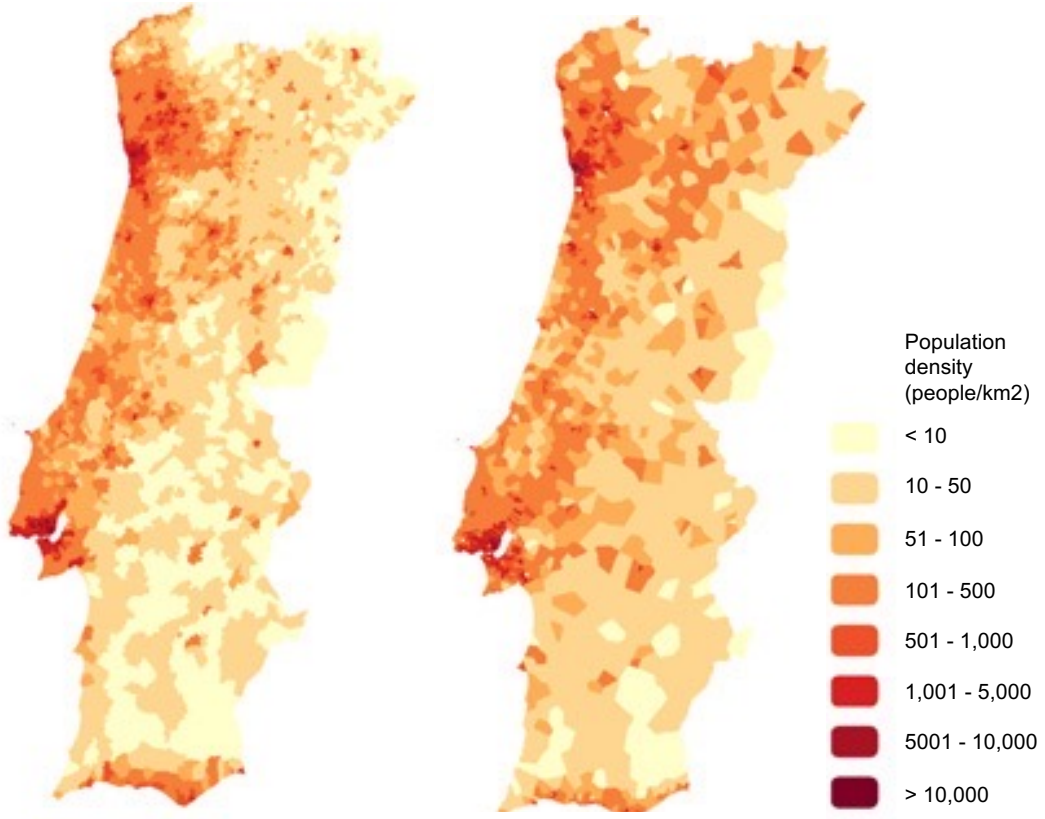
Based on validation study where the results of different home detection algorithms were compared to census counts at cell tower level

Validation of home detection

Validation through:

1. **Aggregate comparison**

Portugal



Census MPD

[Dynamic population mapping using mobile phone data](). Deville et al (2014)

R = 0.89 for municipality level

Lisbon

Validation of home detection

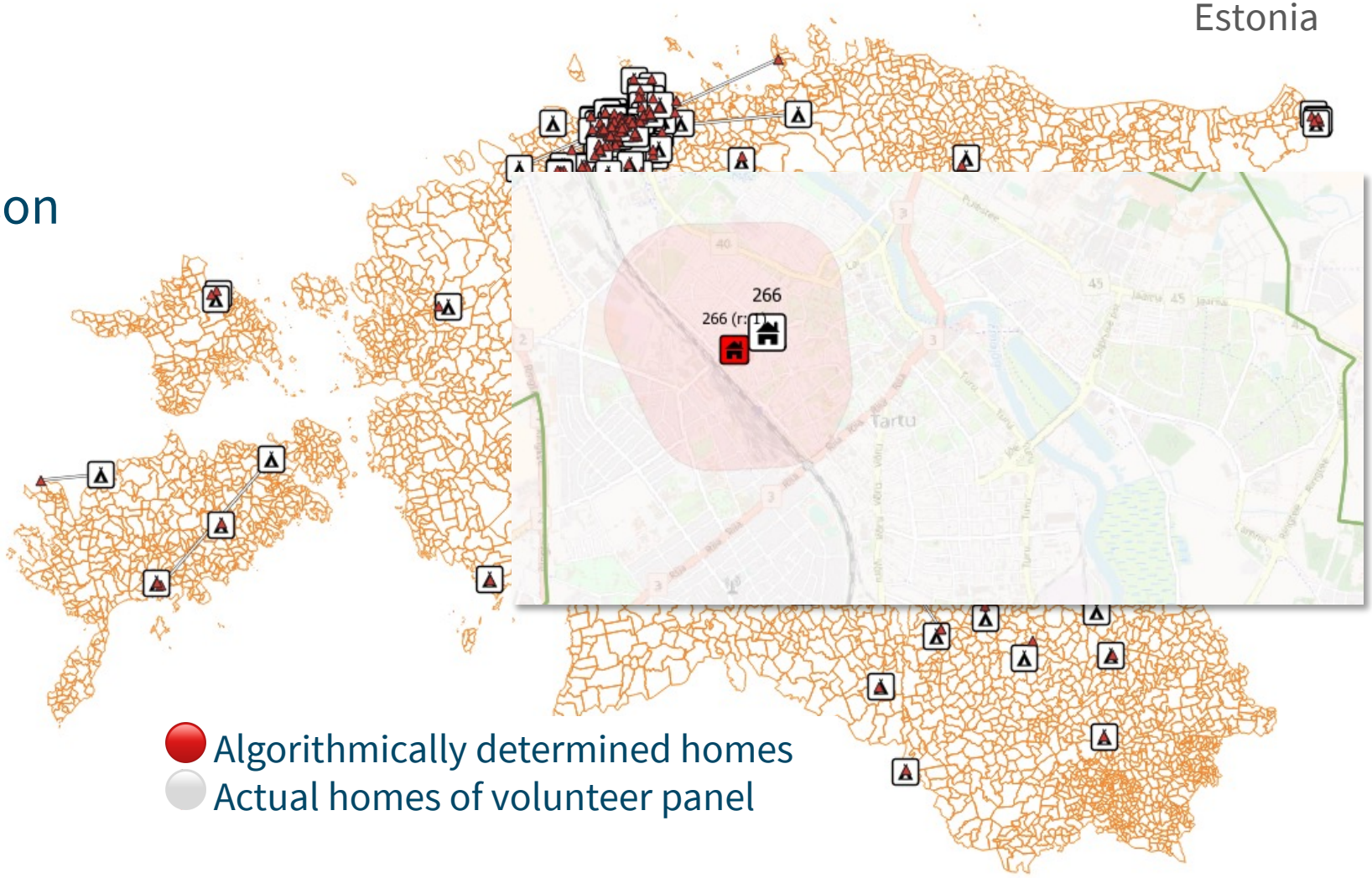
Validation through:

1. Aggregate comparison
2. **Validation panels**

* Ahas et al (2010) Using Mobile Positioning Data to Model Locations Meaningful to Users of Mobile Phones

** Based on internal validation study done with Statistics Estonia, where comparison was done between the MPD anchors calculated by Positium, population registry data and volunteers' true home address

R=0.97 for county level
R=0.82 for polygon



Thank You!

Questions?

siim.esko@positium.com

positium

@Positium