A sensitivity analysis concerning spatial density estimation with MNO data

Authors: Marco Ramljak[1], Fabio Ricciato[2], Peter Lugtig[3]

(1) Corresponding author (m.l.ramljak@students.uu.nl), Utrecht University, (2) Eurostat, (3) Utrecht University

Repository and references

Research Intent

Mobile Network Operator (MNO) data offer a rich source for estimating the spatial distribution of mobile phones at a given point in time. We compare two kinds of estimation strategies: (1) the traditional **Voronoi tessellation and variants thereof**, which exclusively assume that every mobile phone connects to its nearest cell (deterministic), (2) recently introduced estimators that are based on **radio propagation modeling (RPM)** through additional cellular network information (probabilistic). We know that RPM pays off in terms of spatial accuracy if the model information is of perfect quality, however, the question remains how high-quality does the model information need to be?

For this, we have developed two research questions:

- ***RQ1: How robust are estimation strategies to network model mismatching errors (information quality)?***
- ***RQ2: How sensitive are estimation strategies to network characteristics, such as cell density?***

Simulation Study Design

Toyworld Generation

- 100m x 100m grid (1,600 km², right-tailed population distribution, Munich)
- For RQ2, **two synthetic cellular networks** (dense and sparse)
- Mimicking cell counters and developing the generative RPM

Estimation

- **Voronoi estimators:** Tower (VOR_t), antenna offset (VOR_o), barycenter (VOR_b)
- **RPM estimators:** MLE, DF
- For RQ1, **model mismatch** introduced to develop multiple RPMs for estimation

Evaluation

- **Kantorovitch-Wasserstein distance (KWD)** to measure the degree of horizontal spatial error
- Spatial density maps

Key Results

A) Network characteristics

To reserach the effect of network characteristics, we developed two networks of different directional cell-density levels (Fig. 1). Each network leads to a different distribution of overlapping cells per 100m x 100m tile (Fig. 2).

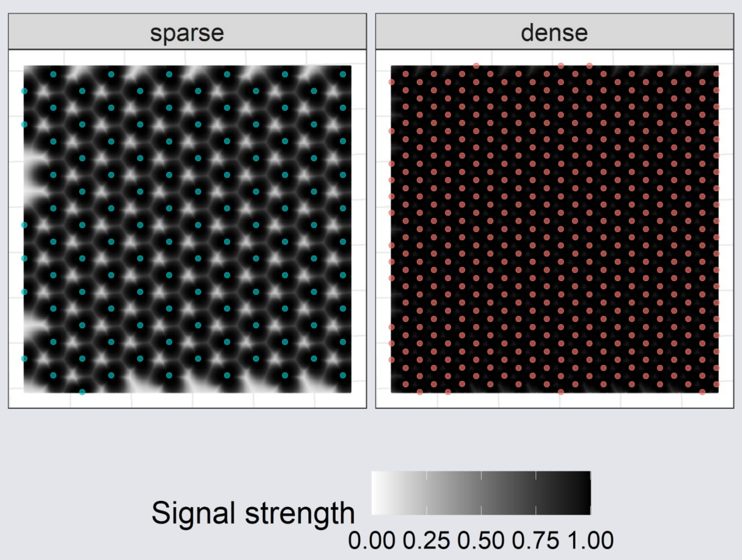

Fig. 1 Two synthetic cellular networks

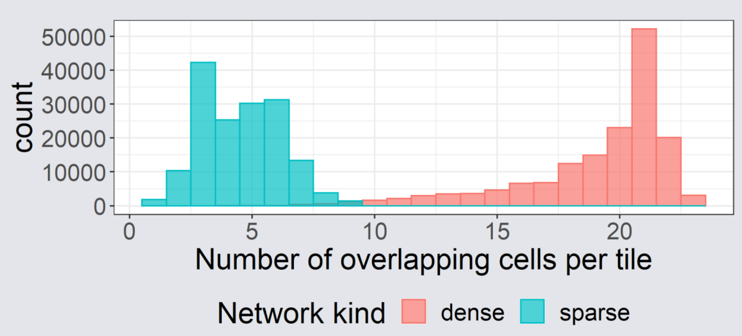

Fig. 2 Cell overlap per tile

B) Model mismatch

To research the effect of network information quality, we use the generative RPM as a basis and introduce spatially sensitve, random noise to create multiple RPMs for estimation. These noised distortions resemble shaded versions of information quality.

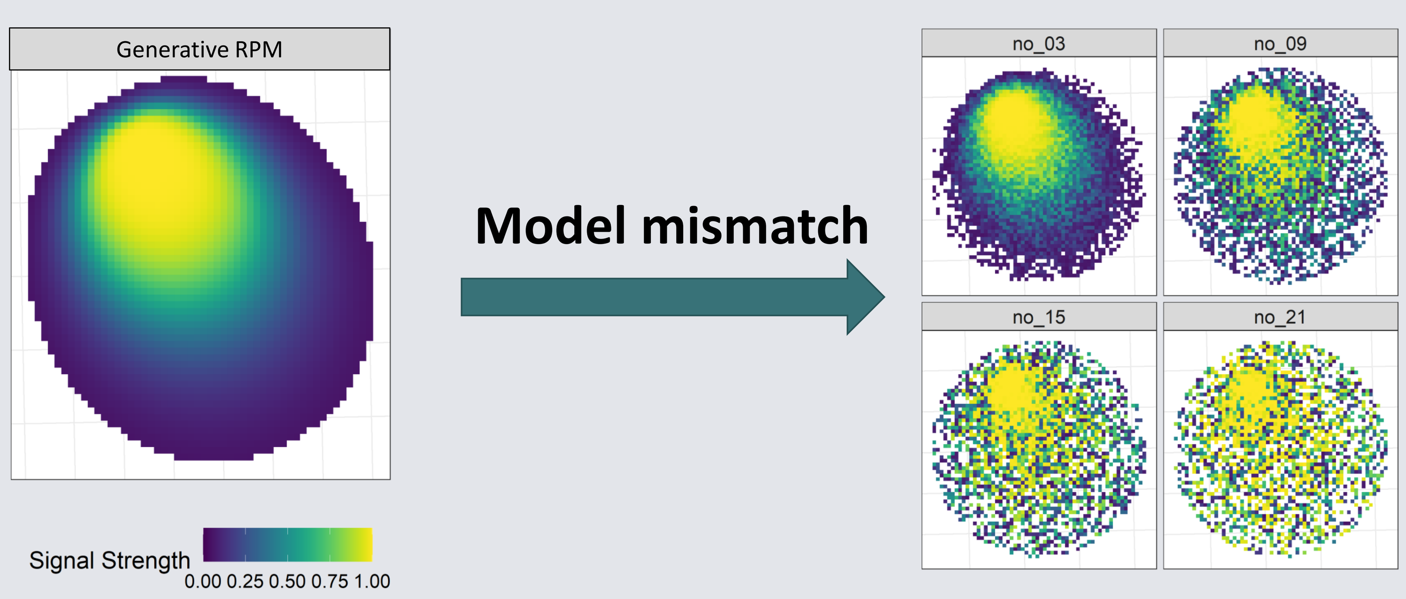

Fig. 3 Model mismatch on exemplary cell profile through random noise

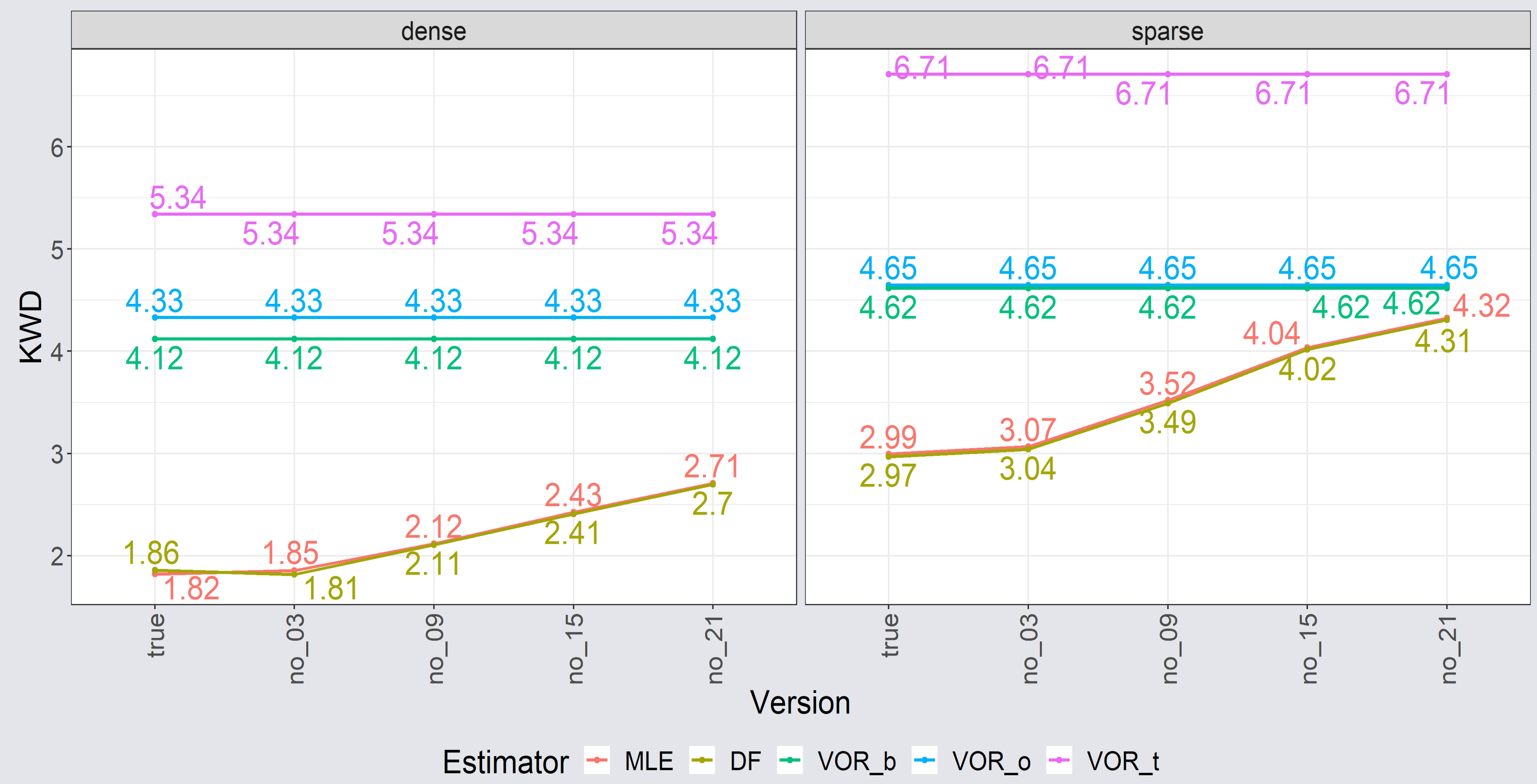

Fig. 4 KWD values for all simulated versions

C) Evaluation

Lower KWD values resemble **higher similarity** to the ground truth density in terms of spatial accuracy. Even in the sparse network scenario the RPM estimators with severely noised RPMs perform much better (Fig. 4). In Fig. 5 one can visually compare the similarity between the ground truth spatial density and the selected MLE estimated spatial densities.

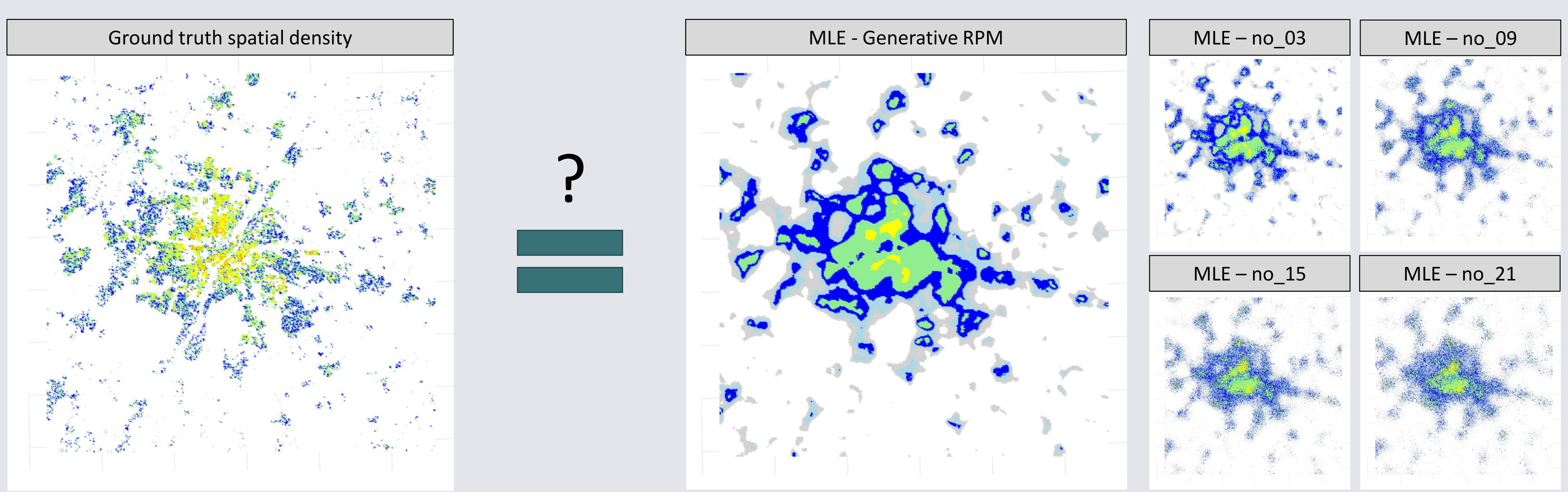

Fig. 5 Selected spatial density maps (comparing ground truth with MLE estimations of the dense network)

Key Takeaways

- ***For RQ1:** Even with severe model mismatch, RPM estimators lead always to spatially more accurate results than any Voronoi estimator.*
- ***For RQ2:** Higher network density offers more information and can therefore lead to spatially more accurate results than sparse networks. RPM estimtators utilize this information better than Voronoi estimators.*

RPM is a complex task, however, in terms of spatial accuracy, it pays off because it allows to model overlapping cells. Our research shows that RPM estimators perform well under different network characteristics and are robust to severe model mismatching errors. Therefore, RPM estimators should be prefered as they utilize various information levels much better than Voronoi estimators. This research project will be extended through more networks (e.g. network layers) and mismatch versions (e.g., quantization, blurring). Further description of the estimators here: arxiv.org/abs/2009.05410