Urban Deprivation in Tanzania: The Bayesian Spatial Bradley—Terry Model

Rowland Seymour[1] | **David Sirl[1]** | **Simon Preston[1]** | **Ian Dryden[1]** | **Madeleine Ellis[2]** | **Bertrand Perrat[2]** | **James Goulding[2]**

[1]School of Mathematical Sciences, University of Nottingham, UK   [2]N/LAB, University of Nottingham, UK

Deprivation in Dar es Salaam, Tanzania

We are interested in the deprivation in different parts of Dar es Salaam, Tanzania. The city of Dar es Salaam is the largest in Tanzania, with over 6 million inhabitants. It is growing at a rapid pace, and is expected to double in size in the next 10 years. Many citizens live in low quality housing, with over 70% living in informal residences.

However, in Tanzania and other less developed countries,

- official statistics are unavailable or unreliable, and
- the population is changing so quickly that available statistics are out of date.

Figure 1: Judges making comparisons

In August 2018, we spent 2 weeks in Dar es Salaam collecting data. We collected over 75,000 comparisons of the 452 areas from over 200 judges. The judges were all citizens of Dar es Salaam, and we found judges via universities in the city and a taxi driver association.

Comparative judgements

We assign to each area what we call a *relative deprivation parameter* $\lambda_i \in \mathbb{R}$ and infer the value of each parameter using a comparative judgement model using the Bradley–Terry model. If areas i and j are compared n_{ij} times, the number of times area i is judged to be more affluent than area j is modelled as

$$Y_{ij} \sim \text{Bin}(n_{ij}, \pi_{ij}),$$

and we assume Y_{ij} are independent. Here the probability π_{ij} that area i is judged to be more affluent than area j depends on the difference in relative deprivation of i and j and is

$$\text{logit}(\pi_{ij}) = \lambda_i - \lambda_j \iff \pi_{ij} = \frac{\exp(\lambda_i)}{\exp(\lambda_i) + \exp(\lambda_j)}$$

The likelihood function for the model is given by

$$\pi(\boldsymbol{y} \mid \lambda_1, \ldots, \lambda_N) = \prod_{i=1}^{N} \prod_{j<i} \binom{n_{ij}}{y_{ij}} \pi_{ij}^{y_{ij}} (1 - \pi_{ij})^{n_{ij} - y_{ij}}.$$

Cities as networks

Urban regions are typically divided into sub-areas for administrative purposes, and these neighbourhoods often provide natural units over which to quantify deprivation. While spatially connected, such areas often vary greatly in size. We model an urban region as a network, whereby these low-level areas are represented as nodes with edges joining neighbouring areas.

Bayesian inference

We use a zero-mean multivariate normal prior distribution for the deprivation parameters $\boldsymbol{\lambda} = \{\lambda_1, \ldots, \lambda_N\}$ subject to the constraint $\mathbf{1}^T \boldsymbol{\lambda} = 0$.

$$(\boldsymbol{\lambda} \mid \mathbf{1}^T \boldsymbol{\lambda} = 0) \sim \text{MVN}\left(\mathbf{0}, \ \Sigma - \Sigma \mathbf{1}(\mathbf{1}^T \Sigma \mathbf{1})^{-1} \mathbf{1}^T \Sigma \right).$$

The covariance matrix is based on the matrix exponential of the adjacency matrix as this measure emphasises connectedness over short distances. Let $\Lambda = e^A$, where A is the network's adjacency matrix, and let D be a diagonal matrix containing the elements on the diagonal of Λ. The covariance matrix is

$$\Sigma = \alpha^2 D^{-\frac{1}{2}} \Lambda D^{-\frac{1}{2}}.$$

Estimated deprivation

We fit the model and estimate the level of deprivation in each area of the city. The results are shown below. There is both a north-south and coastal-interior trend to the level of deprivation. We are also able to estimate the level of deprivation in several slums in the centre of the city.
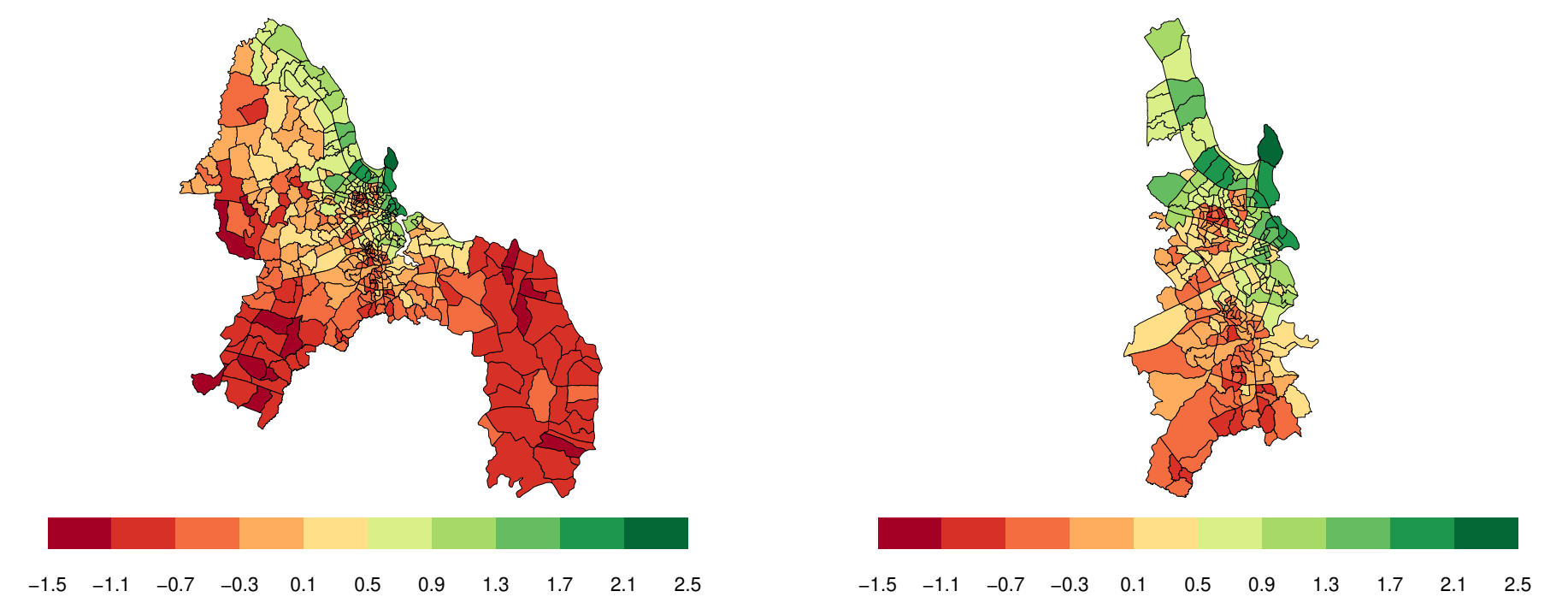

Figure 2: The level of deprivation in each area of Dar es Salaam. The figure on the right shows the centre of the city.

We also analyse how men and women view the city differently. We find that there are parts of the centre of the city that women think are more deprived than the men do.
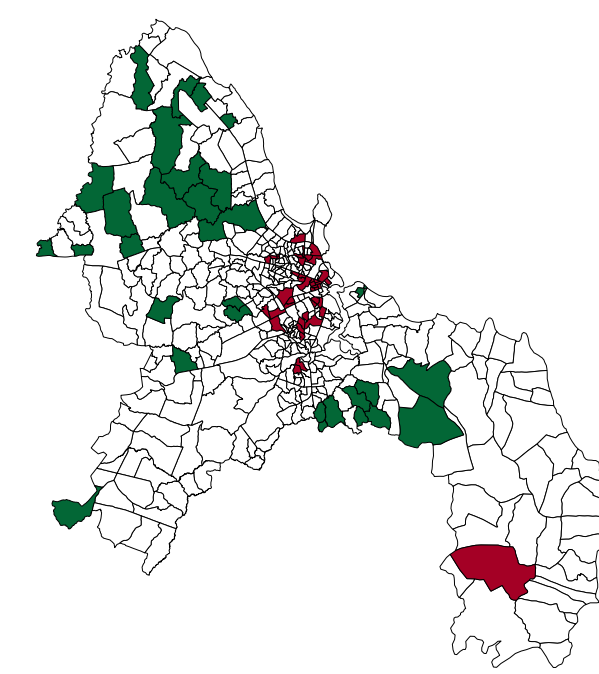

Figure 3: The areas which men and women view differently. The green subwards are viewed as more affluent by women, the red as more deprived.

More information

We have a pre-print available with more details on the modelling
`arXiv:2010.14128`
We have an R package with the data and functions for the analysis
`install.packages('BSBT')`