Retail Planning in Future Cities
A Stochastic Dynamical Singly Constrained Spatial Interaction Model

Mark Girolami

Department of Mathematics, Imperial College London
The Alan Turing Institute
Lloyds Register Foundation Programme on Data-Centric Engineering

https://iconicmath.org/

Imperial College London
Statistical Data Science Conference
03-07-2017

Louis Ellam
ICL & Turing

Mark Girolami
ICL & Turing

Greg Pavliotis
ICL

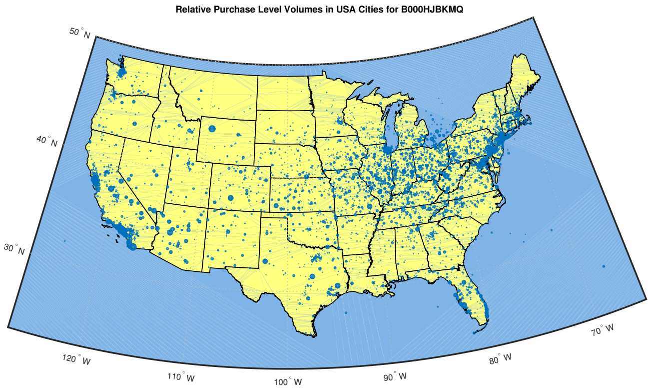
Sir Alan Wilson
UCL & Turing

Question: "What will cities and regions look like in the future?"

Motivation

Question: “What will demand for products be in the future?”

Motivation



Relative Purchase Level Volumes in USA Cities for B000HJBKMQ
50° N
40° N
30° N
120° W
110° W
100° W
90° W
80° W
70° W

Urban Analytics

- How will cities and regions develop as time progresses?
- Observational data from Twitter, Android location tracking, camera monitoring, travel ticketing, demographics, etc.
- Simple statistical summaries and models yielding insights the hallmark of Urban Analytics.
- Mathematical modelling long history in urban and regional analysis and planning.
- Build upon these mathematical formalisms with this new found data.

“We’re uncertain!”

- Urban and regional systems are complex.
- Actions of individuals give rise to an emergent behaviour:
  - Phase transitions;
  - Path dependence; and
  - Multiple equilibria.
- Mechanistic models of complex systems $\implies$ model error.
- Uncertainty should be addressed.

This talk: Improving insights into urban and regional development by addressing uncertainties arising in the modelling process.

The London Retail System

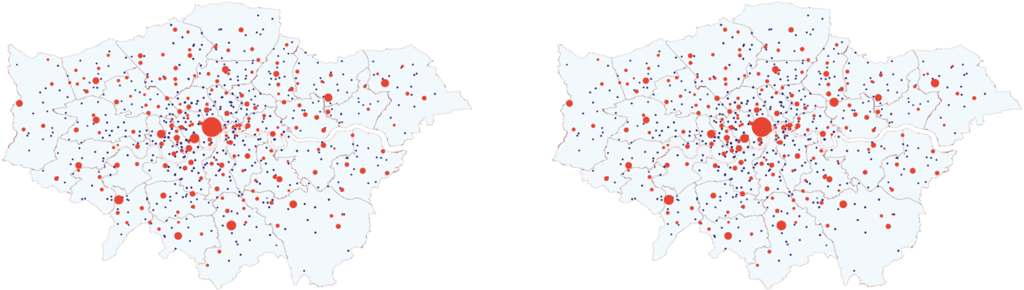


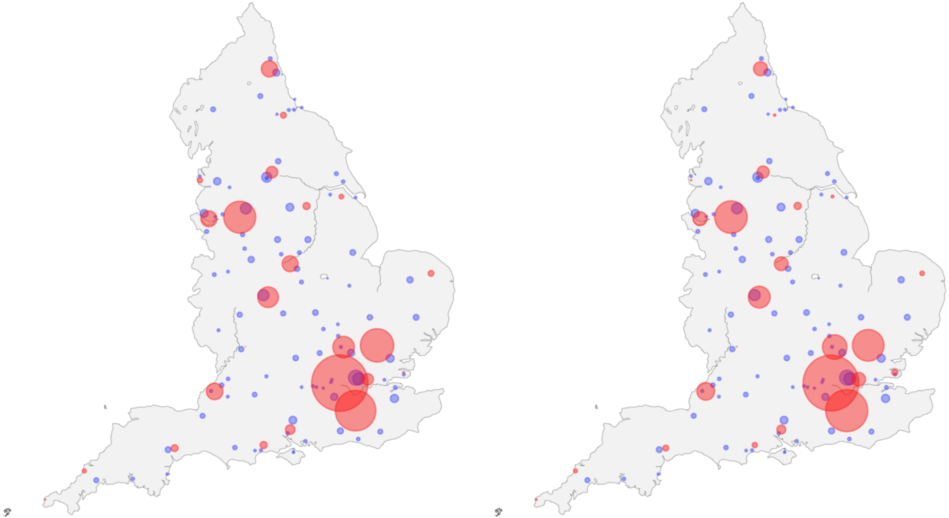
Figure: London retail structure for 2008 (left) and 2012 (right). The locations of retail zones and residential zones are red and blue, respectively. Sizes are in proportion to floorspace and spending power, respectively. $N = 625$ and $M = 201$.

The London Retail System - Schoolboy Statistics

- "Stratford and Shepherds Bush have risen up the rankings as a result of the opening of Westfield Stratford City and Westfield London."
- "Knightsbridge appears to have experienced a significant reduction in town centre floorspace. Ilford and Harrow also experienced declines in town centre floorspace."
- "The West End remains the largest centre in London."
- "Croydon, Kingston, Romford, Canary Wharf, Camden Town, Kings Road East and Angel showed strong growth in total town centre floorspace over the 2007-2012 period."

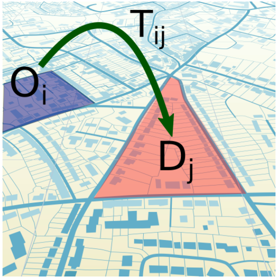
Source: 2013 London Town Centre Health Check Analysis Report.

Airports in England



Urban and Regional Modelling

Spatial Interaction Model



- Assume N origin zones and M destination zones.
- Origin quantities $\{O_i\}_{i=1}^{N}$.
- Destination quantities $\{D_j\}_{j=1}^{M}$.
- $\{T_{ij}\}_{i,j=1}^{N,M}$ denote the flows from zone i to j, respectively.

Assumptions

Flows from origin zones:

$$O_i = \sum_{j=1}^{M} T_{ij}, \quad i = 1, \ldots, N.$$

Flows to destination zones:

$$D_j = \sum_{i=1}^{N} T_{ij}, \quad j = 1, \ldots, M.$$

Known as a singly-constrained or production-constrained system since $\{O_i\}_{i=1}^N$ is fixed and $\{D_j\}_{j=1}^M$ is undetermined.

Assumptions

Fixed transport benefit:

$$\sum_{i=1}^{N} \sum_{j=1}^{M} T_{ij} \ln W_j = X,$$

where W_j is the size, and $X_j := \ln W_j$ is the attractiveness, of zone j.

Fixed transport cost:

$$\sum_{i=1}^{N} \sum_{j=1}^{M} T_{ij} c_{ij} = C,$$

where c_{ij} is the cost of transporting a unit from zone i to zone j.

A Singly-Constrained Spatial Interaction Model

Maximization of an entropy function.

The destination flows are given by

$$D_j = \sum_{i=1}^{N} O_i \frac{\exp\left(\alpha \ln W_j - \beta c_{ij}\right)}{\sum_{k=1}^{M} \exp\left(\alpha \ln W_k - \beta c_{ik}\right)}.$$

- $X_j := \ln W_j$ is the attractiveness of j.
- c_{ij} is the inconvenience of transporting from zone i to j.
- α is the attractiveness scaling parameter.
- β is the cost scaling parameter.
- $\alpha X_j - \beta c_{ij}$ is the net utility from transporting from zone i to j.

A Dynamical Urban and Regional Model

Harris and Wilson Model

The sizes of the destination zones satisfy the (modified) competitive Lotka-Volterra equations

$$\frac{dW_j}{dt} = \epsilon\Big(D_j - \kappa W_j\Big)W_j, \quad j = 1, \dots, M,$$

for some initial condition $W(0) = w$.

Definitions:

- $\kappa > 0$ is the cost rate per unit size; and
- $\epsilon > 0$ is the responsiveness parameter.

Stochastic Urban and Regional Modelling

Potential Energy Function

- In what proceeds, it is more natural to work in terms of attractiveness $X_j := \ln W_j$.
- The 'energy' of a configuration $X = (X_1, \dots, X_M)^T$ is given by an interaction potential $U(X)$.
- The 'force' acting on each site is given by a conservative vector field $-\nabla U(X)$.

Interaction Potential

$$\gamma^{-1} U(X) = \underbrace{\kappa \sum_{j=1}^{M} \exp(X_j)}_{\text{Running cost}} - \underbrace{\alpha^{-1} \sum_{i=1}^{N} O_i \ln \sum_{j=1}^{M} \exp(\alpha X_j - \beta c_{ij})}_{\text{Net utility}} - \underbrace{\delta \sum_{j=1}^{M} X_j}_{\text{Economic Stimulus}} .$$

More General Modelling Requirements

We show that:

1. There **exists a potential function** $U \in C^2(\mathbb{R}^M, \mathbb{R})$ such that

$$-\gamma^{-1}\nabla U(X) = \sum_j R_j(X) - \kappa \exp(X_j) + \delta_j(X).$$

2. The drift coefficients $-\nabla U(X)$ are **locally Lipshitz** and satisfy the **one-sided growth** condition in that there exist $K_1, K_2 > 0$ such that

$$-\nabla U(X) \cdot X \leq K_1 + K_2 \|X\|^2, \quad \forall X \in \mathbb{R}^M.$$

3. U satisfies the **super-linear growth** condition in that there exists $K_1, K_2 > 0$ such that

$$U(X) \geq K_1 + K_2 \|X\|, \quad \forall X \in \mathbb{R}^M.$$

4. The initial condition $X(0) = x$ has **finite variance** $\mathbb{E}\|x\|^2 < \infty$.

Key Results under the Model Structure

Urban dynamics, in terms of attractiveness, satisfies the overdamped Langevin dynamics

$$\frac{dX}{dt} = -\nabla U(X) + \sqrt{2\gamma^{-1}} \circ \frac{dB}{dt}, \quad X(0) = x.$$

The SDE has a unique solution $X \in C([0, T]; \mathbb{R}^M)$ that does not explode on $[0, T]$, almost surely.

Since $\delta_j = \delta > 0$ and $U \in C^2$ with a growth condition, $X(t)$ is ergodic with respect to a well-defined Boltzmann distribution, whose density w.r.t. Lesbgue Measure is given by

$$\pi_\infty(X) = \frac{1}{Z} \exp\big(-\gamma U(X)\big), \quad Z := \int_{-\infty}^{\infty} \exp\big(-\gamma U(X)\big) dX < \infty.$$

A Stochastic Formulation of the Spatial Interaction Model

- With our specification of $U(X)$, overdamped Langevin dynamics give the Harris and Wilson model, plus multiplicative uncertainty.

Stochastic Urban Retail Model

Floorspace dynamics is a stochastic process that satisfies the following Stratonovich SDE.

$$\frac{dW_j}{dt} = \epsilon W_j \Big(D_j - \kappa W_j + \delta \Big) + \sigma W_j \circ \frac{dB_j}{dt},$$

where $\big(B_1, \ldots, B_M \big)^T$ is standard M-dimensional Brownian motion and $\sigma = \sqrt{2/\epsilon\gamma} > 0$ is the volatility parameter.

- Fluctuations (missing dynamics) are modelled as noise that we interpret in the Stratonovich sense.
- The extra parameter (or function) δ to represents local economic stimulus to prevent zones from collapsing (needed for ergodicity).

Metropolised Time-Stepping methods

Metropolis-Adjusted Langevin Truncated Algorithm (MALTA)

Metropolis Hastings algorithm with an Euler-Maruyama proposal, with truncated drift:

$$\hat{X}^* := \hat{X}_n - \frac{\tau \nabla U(\hat{X}_n)}{1 \vee \tau \|\nabla U(\hat{X}_n)\|} + \sqrt{2\tau\gamma^{-1}}\xi_n, \quad \xi_n \sim \mathcal{N}(0, I).$$

- Truncated drift:
  - $\|\nabla U(X)\| \leq \tau$ gives an Euler-Maruyama proposal; and
  - $\|\nabla U(X)\| > \tau$ gives an Euler-Maruyama proposal, but with unit drift.
- Resulting Markov chain is geometrically ergodic with respect to π_∞.
- Converges strongly to the Markov process $\{X(t) : t \in [0, T]\}$ on finite time intervals.

$M = 2$ and $N = 20$:

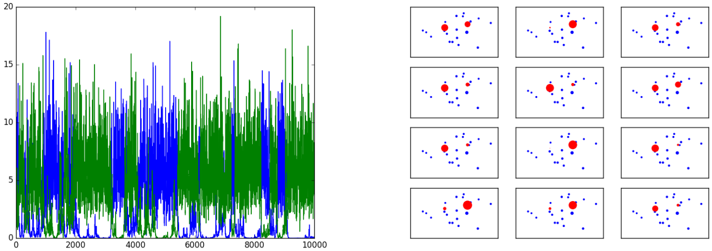


Figure: Sample path of prototype example with parameter values $\alpha = 1.8, \beta = 1, \gamma = 1, \delta = 0.5$ and $\kappa = 1$. Under this parameter regime the two sites do not coexist and there are frequent phase transitions.

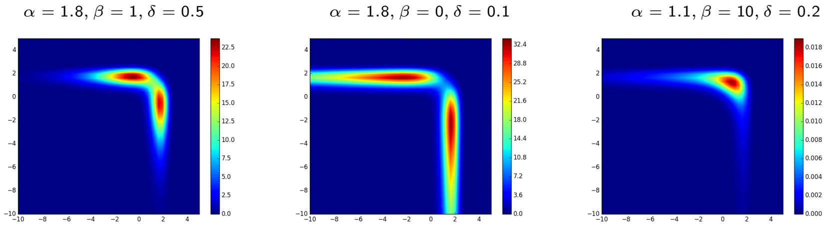
The Boltzmann Distribution - A Prior Measure on Retail Configurations

Insights

- Probability distribution that encodes our knowledge of spatial interactions.
- Configurations are confined to a metastable state (path dependence).
- As $t \to \infty$, system moves towards lowest energy configurations.

Sampling Challenges

- π_∞ is high-dimensional, anisotropic, multi-modal and involves rare events.



Sampling from the Boltzmann Distribution

A 'Simplified' Initial Distribution

Define the density $\pi_0(X) \propto exp(-\gamma_0 U_0(X))$ with

$$U_0(X) = \sum_{j=1}^{M} \left\{ \underbrace{\kappa \exp(X_j)}_{\text{Running cost}} - \underbrace{\left(\delta + \frac{1}{M} \sum_{j=1}^{M} O_i \right) X_j}_{\text{Equal net utility}} \right\},$$

for some $\gamma_0 < \gamma$. We can sample from π_0 exactly and easily.

Bridging Distributions

Define a temperature schedule $0 = t_0 < t_1 < \cdots < t_T = 1$ with bridging distributions

$$\pi_j = \frac{1}{Z_j} \pi_0^{1-t_j} \pi_\infty^{t_j}, \quad \text{where} \quad Z_j := \int \pi_j dX, \quad \text{for} \quad j = 1, \ldots, T.$$

Bridging distributions

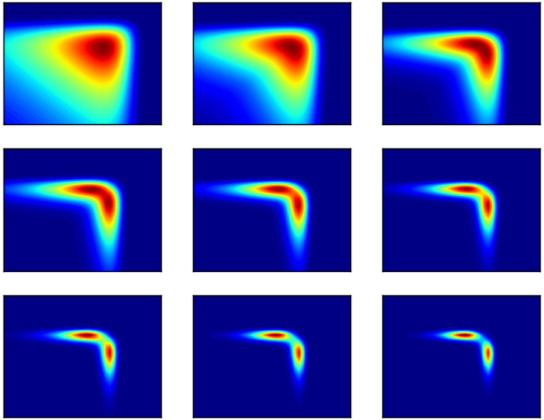


Figure: $\{\pi_j\}_{j=1}^{T=9}$ from left to right, then top to bottom.

The London Retail System: Invariant Distribution

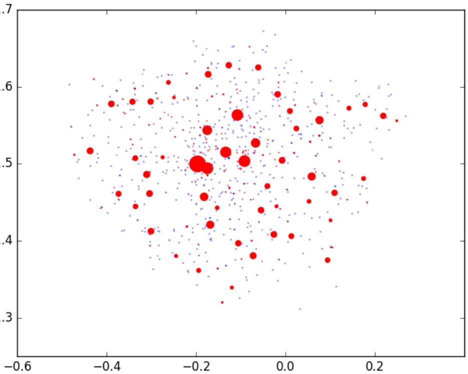


Figure: A sample from the invariant distribution showing behaviour in the limit $t \to \infty$ (most stable configurations). The largest zone is Kensington with 12% of total floorspace and the West End continues to be significant with 5% of total floorspace.

A Statistical Model of Urban Retail Structure

Hierarchical Modelling

Given observation data $Y = (Y_1, \dots, Y_M)$, of log-sizes, infer the parameter values $\theta = (\alpha, \beta)^T \in \mathbb{R}_+^2$ and corresponding latent variables $X \in \mathbb{R}^M$.

Assumption (Data generating process)

Assume that each observation $Y_1, \dots, Y_M$ is an independent and identical realization of the following hierarchical model:

$$Y_1, \dots, Y_M | X, \sigma \sim \mathcal{N}(X, \sigma^2 I),$$

$$X|\theta \sim \pi(X|\theta) \propto \exp(-U(X;\theta)),$$

$$\theta \sim \pi(\theta).$$

Joint Posterior Distribution

The joint posterior is given by

$$\pi(X,\theta|Y) = \frac{\pi(Y|X,\theta)\pi(X,\theta)}{\pi(Y)}, \quad \pi(Y) = \int \pi(Y|X,\theta)\pi(X,\theta)dXd\theta.$$

We have a hierarchical prior given by

$$\pi(X,\theta) = \frac{\pi(\theta)\exp(-U(X;\theta))}{Z(\theta)}, \quad Z(\theta) = \int \exp(-U(X;\theta))dX.$$

The joint posterior is doubly-intractable

$$\pi(X,\theta|Y) = \frac{\pi(Y|X,\theta)\exp(-U(X;\theta))\pi(\theta)}{\pi(Y)Z(\theta)}.$$

Russian Roulette

Figure: On Russian Roulette Estimates for Bayesian Inference with Doubly-Intractable Likelihoods. AM Lyne, M Girolami, Y Atchade, H Strathmann, D Simpson. Statistical Science, 30 (4), 443-467

Metropolis-within-Gibbs with Block Updates

- We are interested in low-order summary statistics of the form

$$\mathbb{E}_{X,\theta|Y}[h(X,\theta)] = \int h(X,\theta)\pi(X,\theta|Y)dXd\theta.$$

- X and θ are highly coupled, so we use Metropolis-within-Gibbs with block updates.
- X-update. Propose $X' \sim Q_X$ and accept with probability

$$\min\left\{1, \frac{\pi(Y|X',\theta)\,\exp(-U(X';\theta))\,q(X|X')}{\pi(Y|X,\theta)\,\exp(-U(X;\theta))\,q(X'|X)}\right\}.$$

- θ-update. Propose $\theta' \sim Q_\theta$ and accept with probability

$$\min\left\{1, \frac{\pi(Y|X,\theta')\,Z(\theta)\,\exp(-U(X;\theta'))\,q(\theta|\theta')}{\pi(Y|X,\theta)\,Z(\theta')\,\exp(-U(X;\theta))\,q(\theta'|\theta)}\right\}.$$

- Unfortunately the ratio $Z(\theta)/Z(\theta')$ ratio is intractable!

Unbiased Estimates of the Partition Function

- We can use the Pseudo-Marginal MCMC framework if we have an unbiased, positive estimate of $\pi(X|\theta)$, denoted $\hat{\pi}(X|\theta, u)$, satisfying

$$\pi(X|\theta) = \int \hat{\pi}(X|\theta, u)\pi(u|\theta)du.$$

- The Forward Coupling estimator (FCE)[1] gives an unbiased estimate of $1/Z$:

$$E[S] = 1/Z.$$

- The idea is to find two sequences of consistent estimators $\{\mathcal{V}^{(i)}\}$, $\{\tilde{\mathcal{V}}^{(i)}\}$, each with the same distribution, such that $\mathcal{V}^{(i)}$ and $\tilde{\mathcal{V}}^{(i-1)}$ are"coupled".

[1]Markov Chain Truncation for Doubly-Intractable Inference, C. Wei, and I. Murray, AISTATS (2016)

Unbiased Estimates of the Partition Function

- Requires N estimates of $1/Z$ using path sampling e.g. annealed importance sampling or thermodynamic integration, for a random stopping time N.
- Coupling between $\mathcal{V}^{(i)}$ and $\tilde{\mathcal{V}}^{(i-1)}$ is introduced with a Markov chain that shares random numbers.
- Variance reduction technique.
- Then the unbiased estimate is given by

$$S := \mathcal{V}^{(0)} + \sum_{i=1}^{N} \frac{\mathcal{V}^{(i)} - \tilde{\mathcal{V}}^{(i-1)}}{Pr(N \geq i)}.$$

The Signed Measure Problem

- S can be negative when $\mathcal{V}^{(i)} < \tilde{\mathcal{V}}^{(i-1)}$ for many i. This is known as the 'sign problem'.
- Rejecting when S is negative would introduce a bias.
- Instead, we note that

$$\mathbb{E}[h(X,\theta)] = \frac{1}{\pi(Y)} \int h(X,\theta)\pi(Y|X,\theta)\hat{\pi}(X|\theta,u)\pi(\theta)\pi(u)dud\theta dX,$$

$$= \frac{\int h(X,\theta)\sigma(X|\theta,u)\check{\pi}(X,\theta,u|Y)dud\theta dX}{\int \sigma(X|\theta,u)\check{\pi}(X,\theta,u|Y)dud\theta dX},$$

where σ is the sign function and we have defined

$$\check{\pi}(X,\theta,u|Y) = \frac{\pi(Y|X,\theta)|\hat{\pi}(X|\theta,u)|\pi(\theta)\pi(u)}{\int \pi(Y|X,\theta)|\hat{\pi}(X|\theta,u)|\pi(\theta)\pi(u)dud\theta dX}.$$

Pseudo-Marginal Markov Chain

- We can sample from $\check{\pi}(X, \theta, u|Y)$ using Metropolis-within-Gibbs with block updates.
- X-update. Propose $X' \sim Q_X$ and accept with probability

$$\min\left\{1, \frac{\pi(Y|X', \theta)\ \exp(-U(X'; \theta))\ q(X|X')}{\pi(Y|X, \theta)\ \exp(-U(X; \theta))\ q(X'|X)}\right\}.$$

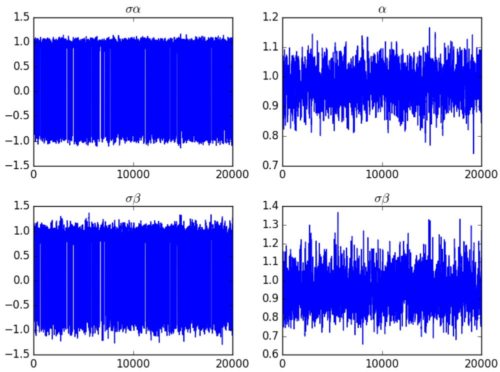
- (θ, u)-update. Propose $(\theta', u') \sim Q_{\theta,u}$ and accept with probability

$$\min\left\{1, \frac{\pi(Y|X, \theta')|S(\theta, u)|\exp(-U(X; \theta'))\pi(\theta')\pi(u')\ q(\theta|\theta')}{\pi(Y|X, \theta)|S(\theta', u')|\exp(-U(X; \theta))\pi(\theta)\pi(u)q(\theta'|\theta)}\right\}.$$

- Posterior expectations are estimated using

$$\mathbb{E}_{X,\theta|Y}[h(X, \theta)] = \lim_{N\to\infty} \frac{\sum_{i=1}^N h(X_i, \theta_i)\sigma(X_i|\theta_i, u_i)}{\sum_{i=1}^N \sigma(X_i|\theta_i, u_i)}.$$

Illustration: Airports in England



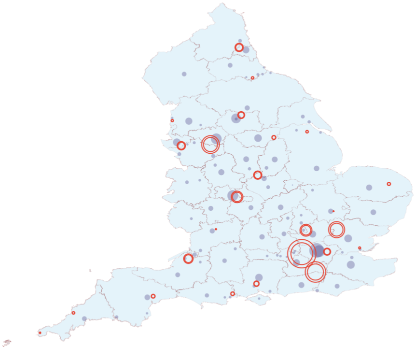
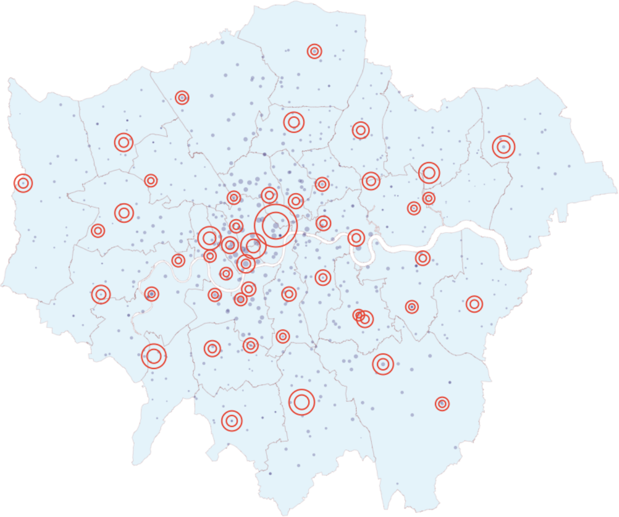


Figure: Left: Trace plots of α and β. 10.1% of the σ values are negative. The posterior mean of α and β are 0.97 and 0.94, respectively, indicating a moderate attraction towards to larger but fairly local airports. Right: 5th and 95th percentiles of the latent states X. There tends to be more variability in larger airports such as Heathrow and Manchester.

Illustration: Retail Centres in London



Conclusions and Outlook

Further work and extensions

- Investigation of new data assimilation methodologies to calibrate models to data available at different scales. For example:
  - Population data;
  - Cost matrix; or
  - Time dependent parameters.
- Deployment of new methodology to a global problem to provide new insights into urban retail structure.
- Improved mechanistic models of urban and regional phenomena.
- Melding of data and models takes us beyond data analytics.

Summary

- We have revisited the Harris and Wilson model to formally account for uncertainties in the modelling process.
- A stochastic extension of the model, under certain conditions, may be expressed as a Langevin diffusion.
- The invariant distribution provides new insights into urban and regional dynamics, and suggests the introduction of a new parameter in the model.
- We presented a Bayesian hierarchical model for urban and regional systems. We have demonstrated our approach using an example of airports in England and retail in London.

Acknowledgment

- EPSRC Programme Grant - EP/P020720/1
- Inference, Computation, and Numerics for Insights into Cities - ICONIC
- `https://iconicmath.org/`