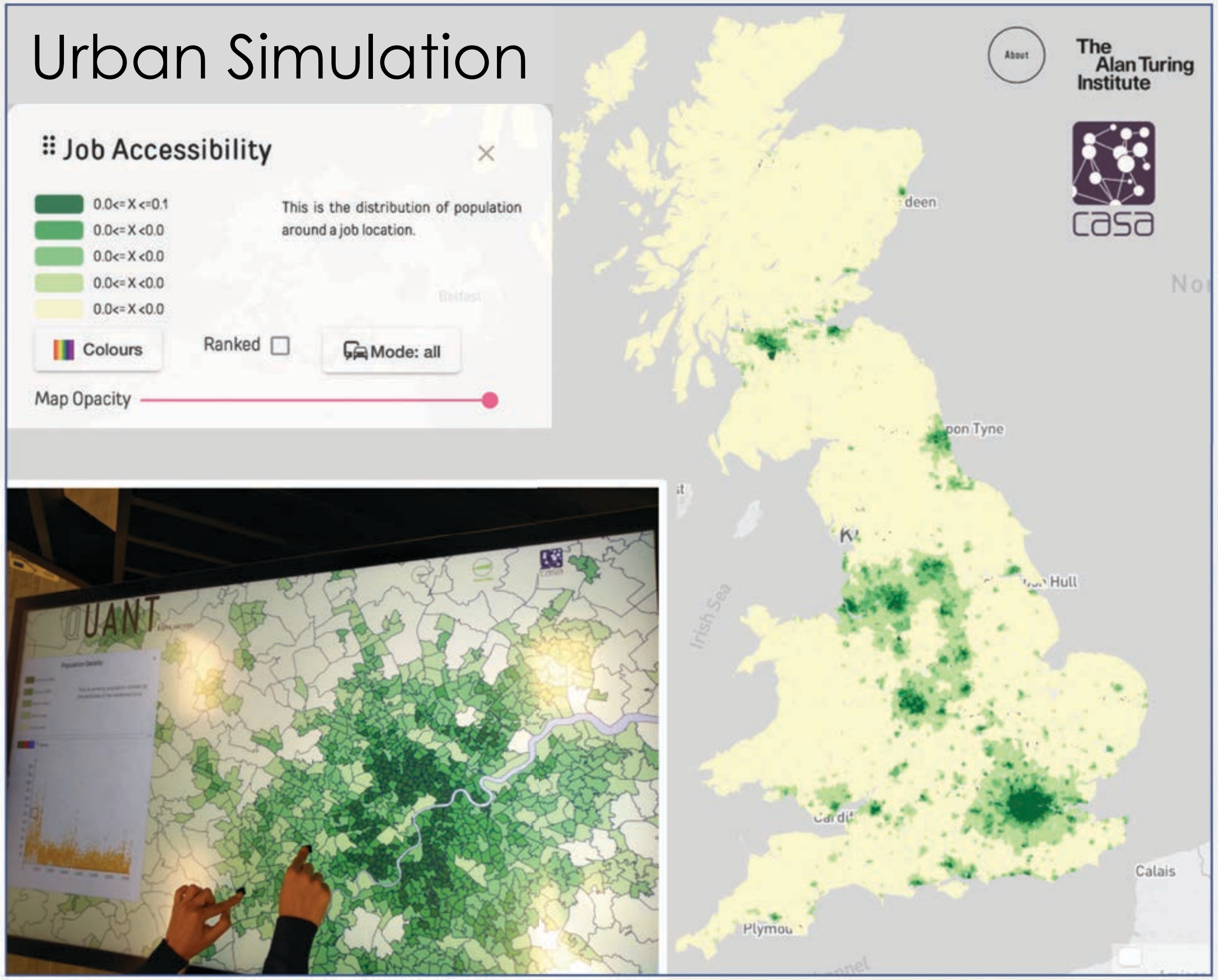
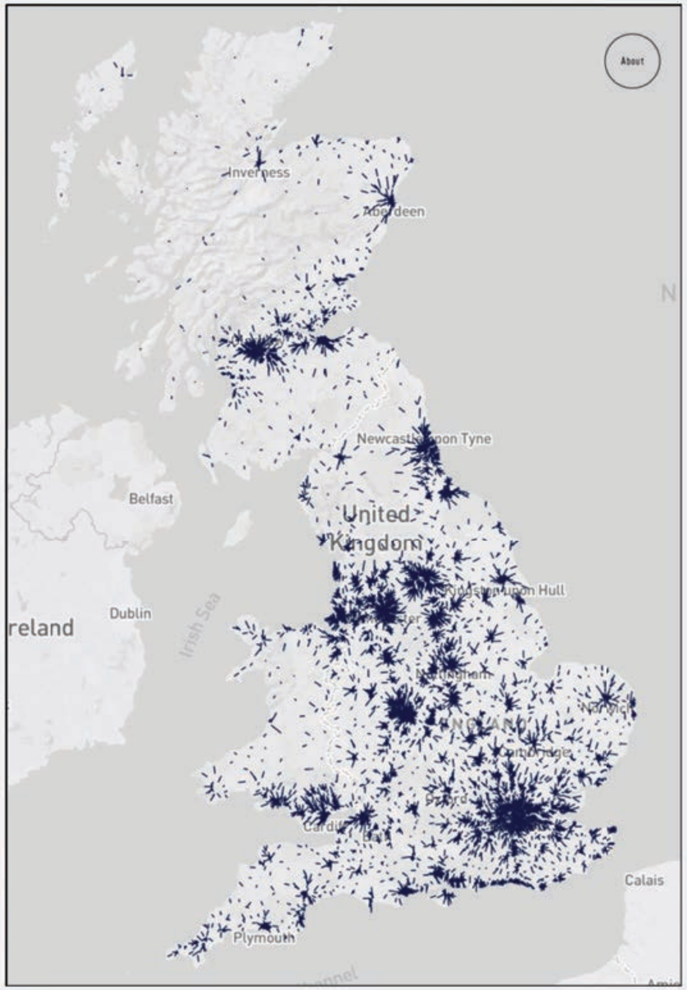
CASA0002: Urban Simulation (2024)





Michael Batty Simulation ● **Elsa Arcaute** Networks

Simulation

Models, Spatial Interaction, LUTI, CA & ABM

Michael Batty

Networks

Elsa Arcaute

Monday, 8 January 2023

m.batty@ucl.ac.uk ● e.arcaute@ucl.ac.uk
X@jmichaelbatty ● **X**@ElsaArcaute

Urban Simulation 1:
Some Preliminaries: What are Models? Relationships to Theory

Michael Batty

m.batty@ucl.ac.uk
@jmichaelbatty

Monday, 9 January 2022

Some Preliminaries: What are Models?

A <u>theory</u> is an *abstraction* of some phenomenon, usually the phenomenon '*real*' but sometimes it is 'imagined' in a form that makes the *simplification* or *abstraction* clear.

A <u>model</u> is a simplification of '*reality*' which takes the theoretical abstractions and puts it into a form that we can manipulate. <u>Simulation</u> is often used to characterise this process of implementation, usually being digital.

In everything we do, we theorise, and more and more frequently **we build models** to demonstrate theory. **This entire course, even the MSc, is about models**

This is all fairly obvious – but the focus on theory is important because theory can be implicit as well as explicit. In fact in our growing quest to describe the world through models, theory is tending to become part and parcel of models.

The main reason for beginning with theory is that the conventional wisdom of science begins with theory and then *tests* theory against observations – data. It is impossible to approach the world without prior theory without being involved where theory comes from.

Let us assume that whenever we model a phenom-enon we have in mind some theory no matter how implicit. It is sometimes hard to even extract the theory we hold but it always there.

Thus the model- building process is really part and parcel of the scientific process – the scientific method – where the current wisdom is that science tests theory by assembling data about reality which is designed to '*falsify*' the theory. This is scientific method *a la Popper* and it suggests that data or observations is the ultimate arbiter of good theory. In some sense our models are always false.

The method implies that this process of testing takes place in systems which are controllable in some sense, as in an experimental lab. In fact as science has progressed, these conditions appear to be increasingly unlikely. Moreover when we take powerful physical theory out of the lab, it is subject to volatility& can rarely achieve predictive success as in the lab

Moreover when we take powerful physical theory out of the lab, it is subject to volatility& can rarely achieve predictive success as in the lab.

In fact our computer becomes our lab – and in computers like in physical labs, we are able to simulate phenomena in theoretical conditions which do not imply the models are right. In fact, in the 1970s, George Box, the statistician said the following

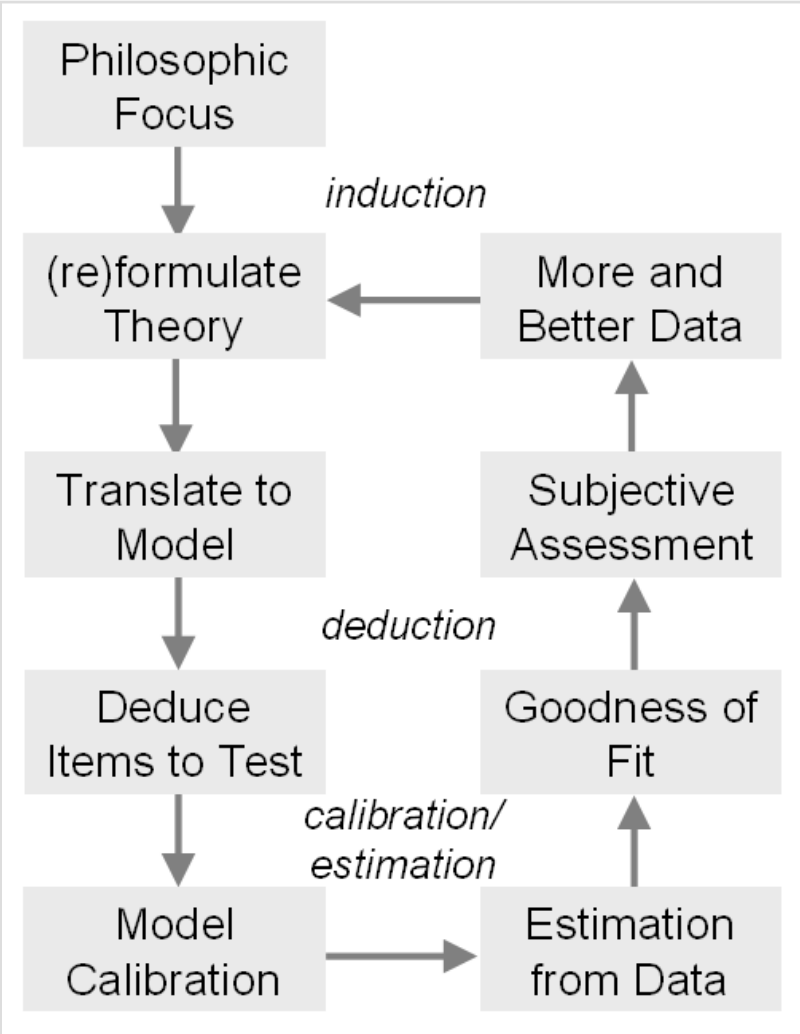
"All models are wrong, but some are useful"

Hence the need for models – for theories in a form other than in the laboratory, where we can perform good testing.

The new form of the laboratory is the computer and instead of experimentation, there is simulation. We could and perhaps we should spend time talking about this issue – for by no means all models are simulation models and all science is not yet based on computers.

But increasingly science is intrinsically about computation – it is informed by digital tools – computer models – and this is changing science itself. I also use the term ' science' advisedly, in its most catholic sense ….. another debate perhaps.

Let me get some more terms out of the way – and to do this here is a simple picture of the scientific method. We follow this in building a model



Let me also say something about

induction

and

deduction

Definitions of Models

There are of course many types of models and although you may think that here we are only going to deal with mathematical or symbolic models, this is not so – in CASA we build material and iconic models as well. Lowry's 1965 paper – "A Short Course in Model Design" paper that I recommend you read, classifies models, and we will draw loosely on his scheme.

Lowry, I. S. (1965) A Short Course in Model Design, **Journal of the American Institute of Planners**, **31**, 158-165.

There seem to be four different but generic ways of abstraction – **iconic**, **analog**, **symbolic** and **logic** but these categories are not mutually exclusive.

A Classification: Icons, Analogues, Symbols

Iconic models are representations that visually convey what the real things looks like – maps are the classic example – these are largely representations although they may have some symbology but they are scaled down versions of the real thing.

Symbolic models represents systems in terms of the way they function, often through time and over space – these models are invariably mathematical.

Analog models are a half way house between iconic and symbolic. The key issue is that they take a representational and/or functional form of one system and apply it to another.

e.g. analogies between physical and human systems – the flow of blood in analogy to hydrodynamics developed for models of the atmosphere, traffic flow as an analogue of an electrical network, and so on.

Logical models are symbolic in a sense but are based on causal connections composed of rules. We can mix, of course, any of these four types.

To this I am going to add Data-driven models. We can look at models existing on a spectrum from the data we collected about a situation that can contain elements of prediction within to fully predictive models that attempt to forecast the future or rather forecast events that have not yet happened or we have not yet observed

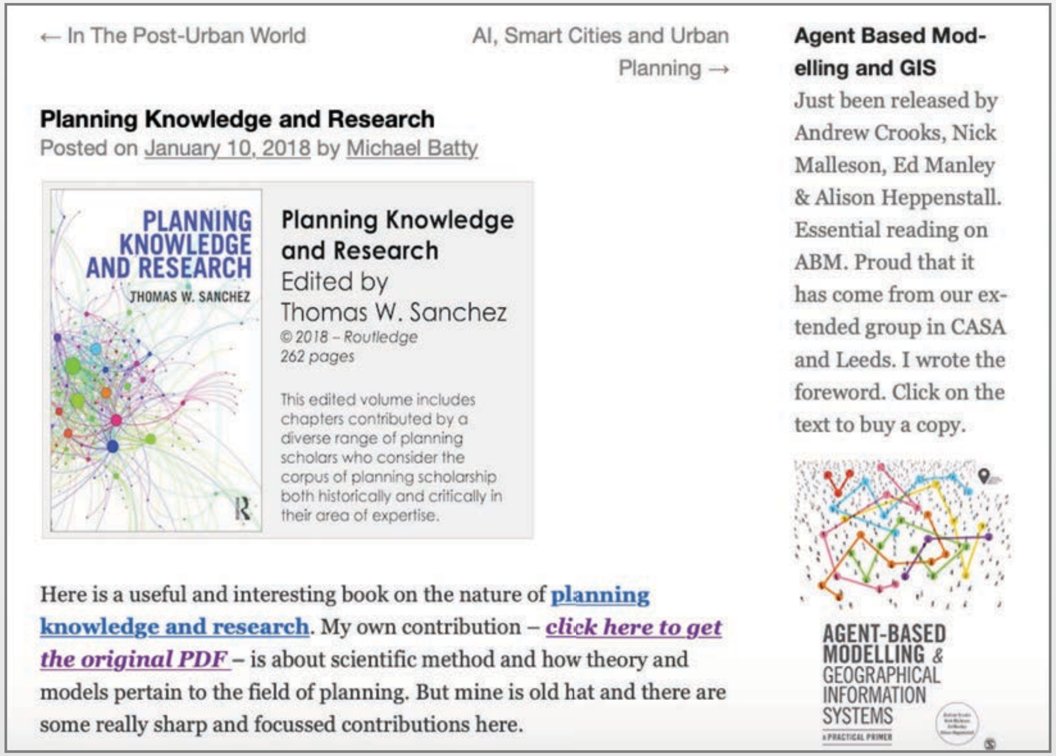
An Emerging Concept: The Digital Twin

As the model gets closer to the real thing, then eventually it merges with the real thing – it becomes the real thing. This of course is impossible for it would no longer be a model. But close to the real thing, we often call this a digital twin. This is an idea that has become significant recently in the last 5 years or so.

The best example is the map – and Lewis Carroll made a parody of this in the late 19^{th} century. I am not going to repeat this but is all about increasing the scale of the map until it covers the whole country. I've written a couple of pieces on this but I will put them on the web site and you can look at them at your leisure but there are links to them below.

Background Reading

There is a paper on science and planning on my blog which is also in the book edited by Sanchez



← In The Post-Urban World

AI, Smart Cities and Urban Planning →

Planning Knowledge and Research

Posted on January 10, 2018 by Michael Batty

PLANNING KNOWLEDGE AND RESEARCH

THOMAS W. SANCHEZ

Planning Knowledge and Research

Edited by Thomas W. Sanchez

© 2018 – Routledge

262 pages

This edited volume includes chapters contributed by a diverse range of planning scholars who consider the corpus of planning scholarship both historically and critically in their area of expertise.

Here is a useful and interesting book on the nature of planning knowledge and research. My own contribution – *click here to get the original PDF* – is about scientific method and how theory and models pertain to the field of planning. But mine is old hat and there are some really sharp and focussed contributions here.

Agent Based Modelling and GIS

Just been released by Andrew Crooks, Nick Malleson, Ed Manley & Alison Heppenstall. Essential reading on ABM. Proud that it has come from our extended group in CASA and Leeds. I wrote the foreword. Click on the text to buy a copy.

AGENT-BASED MODELLING & GEOGRAPHICAL INFORMATION SYSTEMS

A PRACTICAL PRIMER

http://www.spatialcomplexity.info/files/2018/01/Science-in-Planning-Theory-Methods-Models-.pdf

Besides the science and planning paper on my blog, there are some papers worth looking at

Epstein, J. M. (2009) Why Model, **Journal of Artificial Societies and Social Simulation, 1**, no. 4, 12, https://www.jasss.org/11/4/12.html

Lowry, I. S. (1965) A Short Course in Model Design, **Journal of the American Institute of Planners**, **31**, 158-165.

Vanderleeuw, S. E. (2004) Why Model? **Cybernetics and Systems: An International Journal**, **35**, 117-128

Two short pieces on digital twins and what they mean in terms of models

Editorial

B Urban Analytics and City Science

Digital twins

Environment and Planning B: Urban Analytics and City Science
2018, Vol. 45(5) 817–820

DOI: 10.1177/2399808318796416
journals.sagepub.com/home/epb
SAGE

Editorial

B Urban Analytics and City Science

A map is not the territory, or is it?

EPB: Urban Analytics and City Science
2019, Vol. 46(4) 599–602

DOI: 10.1177/2399808319850652
journals.sagepub.com/home/epb
SAGE

The Simplest Models

Changes in Scale and Space: Agglomeration, Density, Gravitation, Rank-Size and Spatial Interaction

Michael Batty

m.batty@ucl.ac.uk

X @jmichaelbatty

Monday, 8 January 2023

Outline

- The Mathematics of Linear and Non-Linear Models
- Gravitation: The Basic Models
- Scale and Agglomeration
- Rank-Size: City Size
- Generalising to Discrete Systems From Continuous
- Reading for these lectures

The Mathematics of Linear and Non-Linear Models

Let me begin by introducing relationships that you are already familiar with – these are simple linear and nonlinear equations we will use extensively for modelling spatial interactions, population densities, etc..

All our models deal with relationships between what we call a ***dependent variable***, something we want to explain and one or more ***independent variables*** that are assumed to explain the variation in the dependent variable.

The dependent variable is sometimes called ***endogenous*** or “coming from within’ and the independent ***exogenous*** meaning “coming from without’

If we have more than one set of relationships and they refer to each other, then these are configured in linked systems that are sometimes called “econometric’ mainly after quantitative economics from which many originate.

Now the simplest of such models are usually linear where the dependent variable increases ***at the same rate*** as the independent variable or variables.

Nonlinear relationships are where the dependent variable increases at ***a lesser rate*** than the independent or ***a greater rate*** than the independent.

If the nonlinear relationship can be transformed into a linear one, then the nonlinear relationship is ***intrinsically linear***. If the relationships cannot be so transformed it is ***intrinsically nonlinear*** – let is look at some of these

The simplest linear relationship is

$$y = a + bx$$

This leads to what is clearly a straight line plot where it is clear that the parameter b is the rate of change which in the calculus is the first derivative which we can compute as

$$\frac{dy}{dx} = b$$

The other parameter in the linear model is a and in terms of the straight line plot a is the intercept and b is the slope. The simplest non-linear relationship is

$$y = ax^b$$

where the rate of change is now a variable where the rate of change is now a variable

$$dy/dx = abx^{b-1}$$

In short it depends on x

Now we say that a relationships is intrinsically linear if we can transform it simply to a linear one and use the simplest methods to estimate it – to fit is as a straight line and then transform it back to the original relationship. The nonlinear relationship above is intrinsically linear – if we take logs it become linear

$$log\, y = log\, (ax^b)$$

And this becomes

$$\log y = \log a + b \log x$$

You can see this is a linear equation and once we have it in this form it is easy to estimate the parameters a and b. Now an intrinsically non-linear relationships is a much more complex expression where a simple transformation does not yield a linear relation.

This is all we need but to make things clearer, we are likely to express these relations with respect to a single dependent variable y and several independent variables $x_{1,}$ $x_{2,}$ $x_{3,}$ ….. and parameters $b_{1,}$ $b_{2,}$ $b_{3,}$ …. with the linear and nonlinear relations as

$$y = a + b_1x_1 + b_2x_2 + b_3x_3 + \cdots$$

$$y = a{x_1}^{b_1}{x_2}^{b_2}{x_3}^{b_3}\ldots.$$

One last complication or rather it is a simplification. We can write these equations using the summation sign as

$$y = a + \sum_{k=1}^{N} b_kx_k = a + b_1x_1 + b_2x_2 + b_3x_3 + \cdots$$

$$y = a\prod_{k=1}^{N} {x_k}^{b_k} = a{x_1}^{b_1}{x_2}^{b_2}{x_3}^{b_3}\ldots$$

So we have linear and nonlinear – as a thought experiment you should now look at these equations and their parameters and see what happens in the simple linear and non linear cases what happens when $0 \leq a \leq +$ and when $- \leq b \leq +$

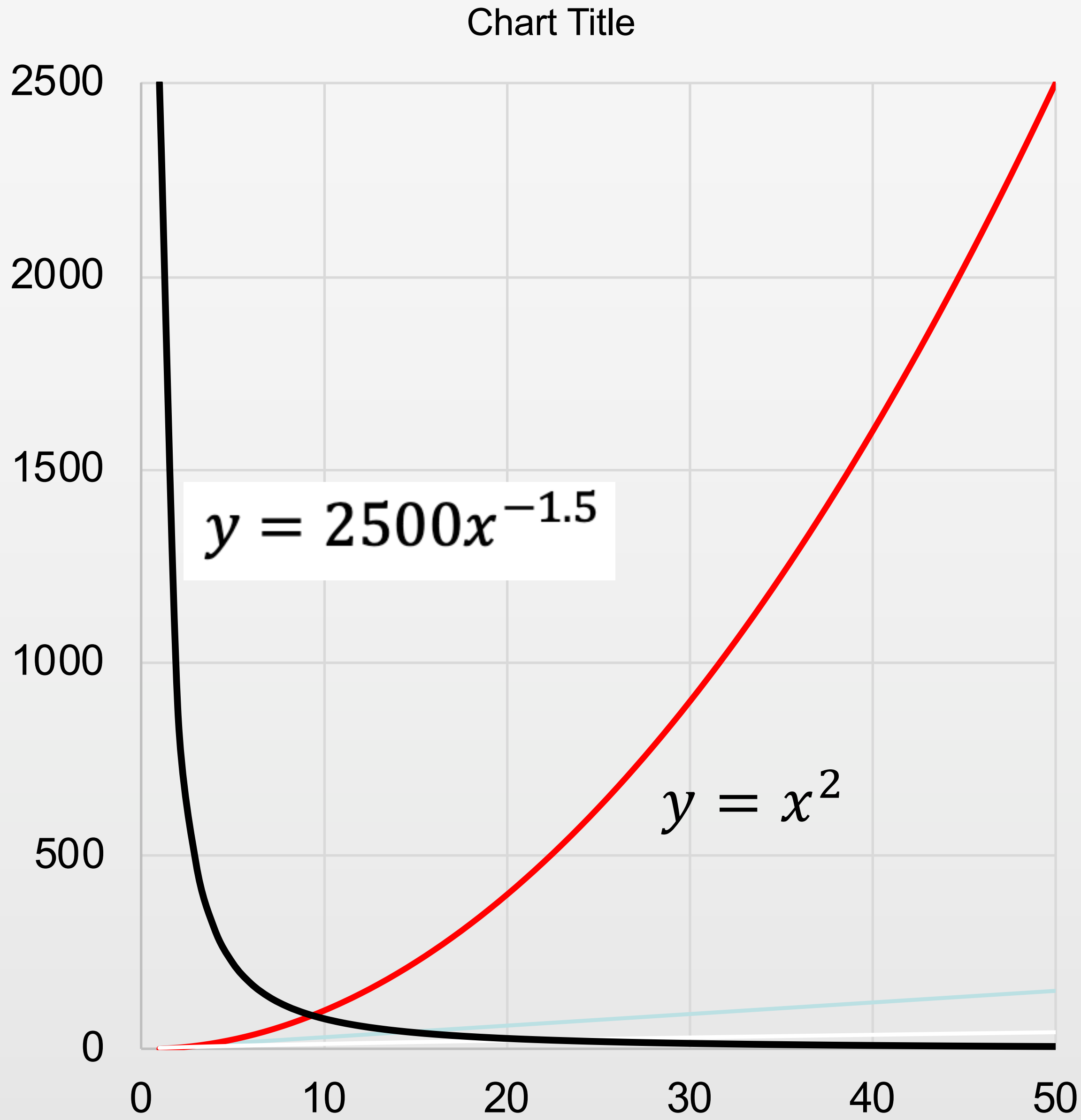
In particular we will be dealing with models that are like

$$y = ax^{-b}$$

where $a > 0$ and $0 \leq b \leq 3$. Note that we are assuming b is positive but when it goes into the equation it is negative. This is the classic inverse power function that we show on the graph as an inverse square law $y = 2000x^{-2}$ where x goes from 1 to say 50

This is the classic inverse power function that we show on the graph as an ***inverse power law*** $y = 2500x^{-1.5}$ where x goes from 1 to say 50

The basic law in physics is the ***inverse square law*** – Newton's Law $y = 2500x^{-2}$



Gravitation: The Basic Models

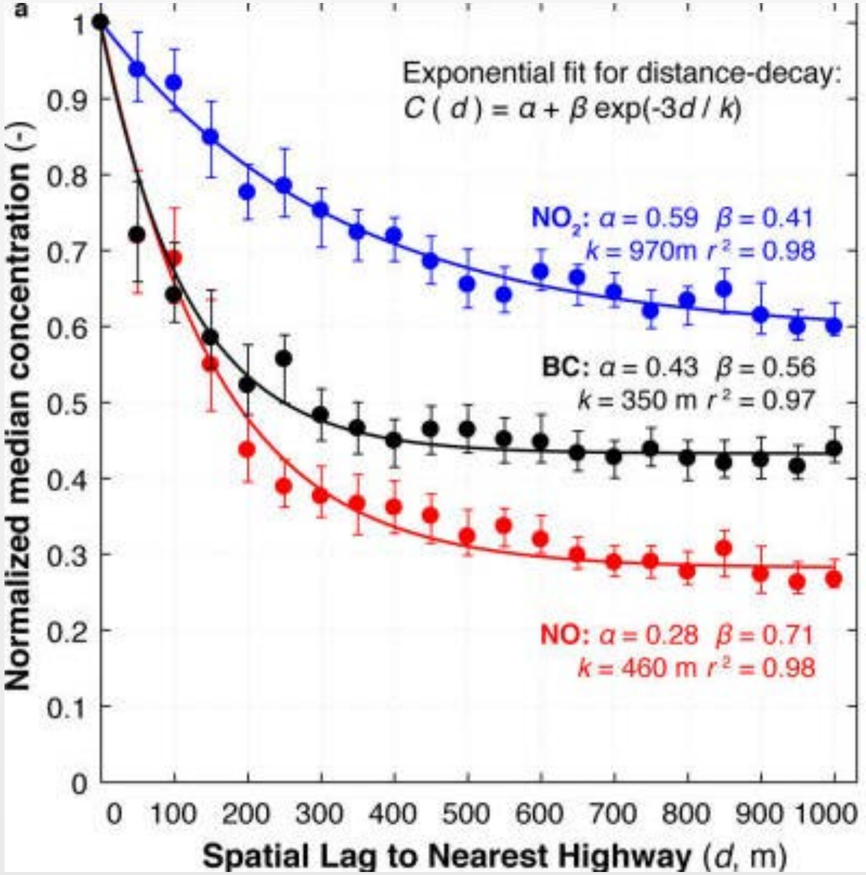
Now power laws occur everywhere in systems where the elements compete with one another. The most basic is that which relates to spatial distributions. In general, all other things being equal (ceteris paribus), the flow or volume of trips $T(d)$, the spatial interaction from a location – sometimes called an origin – to all other locations – destinations – at increasing distances from the origin declines inversely with that distance d. Then we can write this equation as above as

$$T(d) = Kd^{-\beta}$$

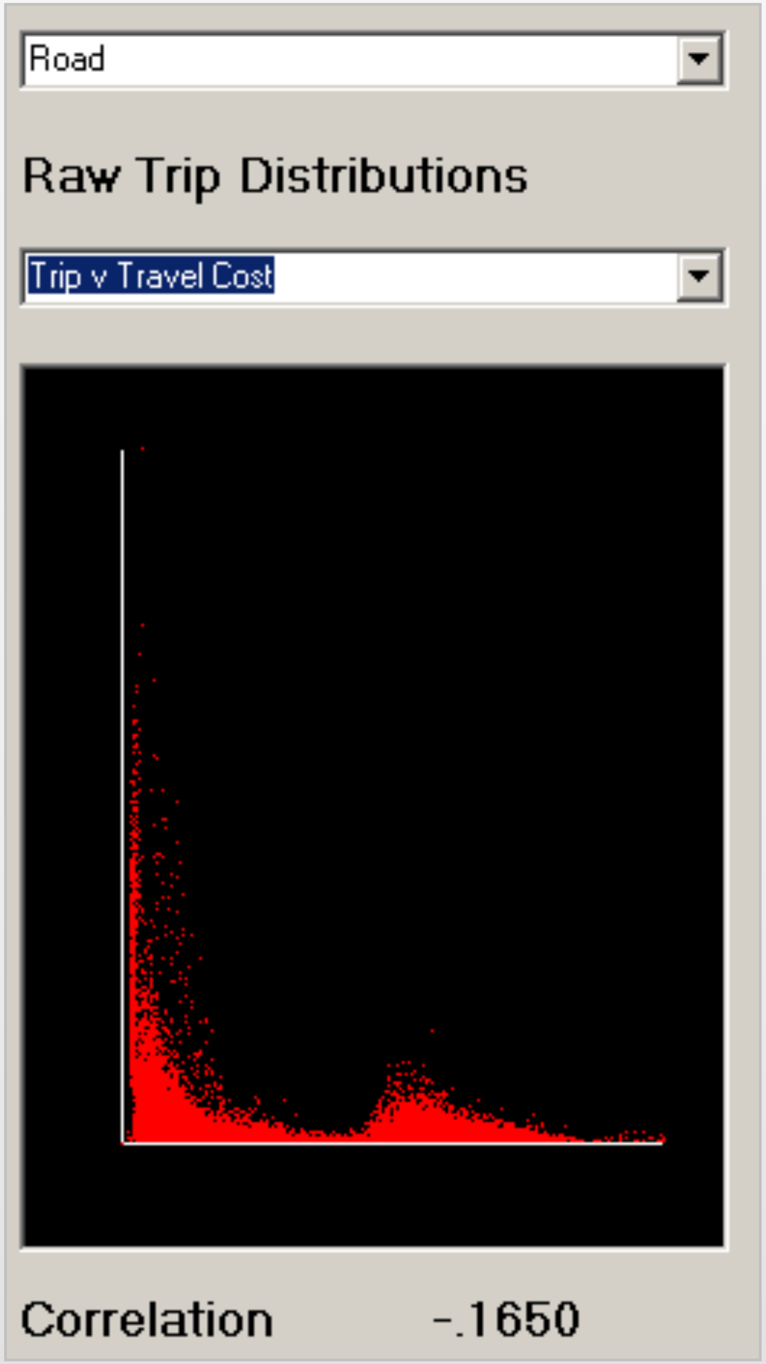
There are hundreds of demonstrations of this so called distance effect - note we can also use proxies for distance like travel cost, travel time and so on – in fact this feature of cities and regions has been called Tobler's Law which states that "everything is related to everything else, but near things are more related than distant." Waldo Tobler said this in a paper in 1970 but it is generic – sometimes it is called the first law of geography. Look at the Wikipedia page for details

https://en.wikipedia.org/wiki/Tobler%27s_first_law_of_geography

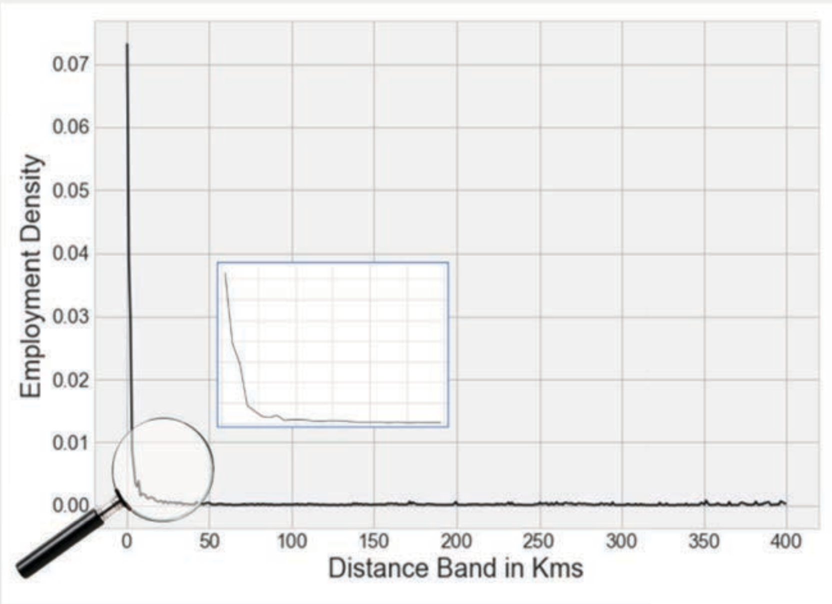
Lets look at some examples

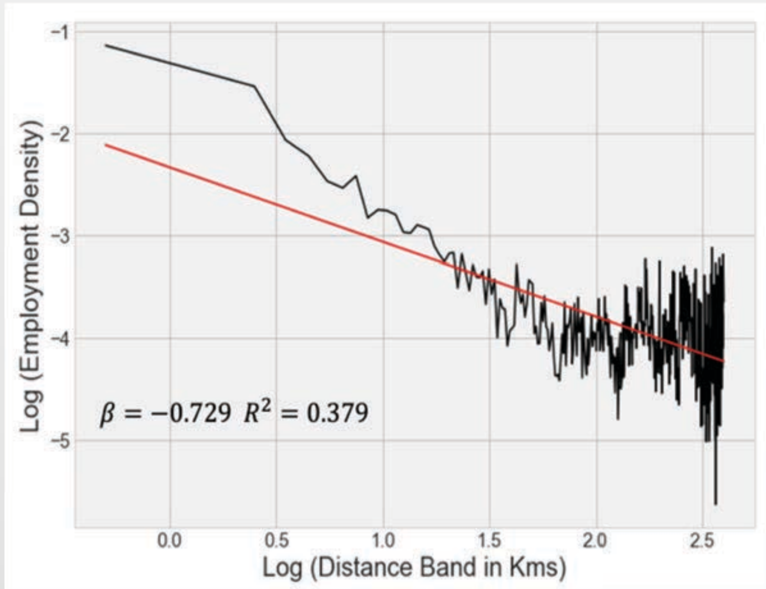


Decline of Trips
with Distance



In fact the inverse power model wrt to a single origin which is the city centre or CBD is in fact the population density model first developed by Clark in 1951 and generalised by him to a negative exponential. What we have done here is to look at trip density as population in successive distance bands of 1 km width from central London – Charing Cross





The Clark model replaces the inverse power with the inverse or negative exponential and the model then becomes

$$\rho(D) = Zexp(-\gamma d)$$

Compare this to

$$T(d) = Kd^{-\beta}$$

Now in all gravity or spatial interaction models that assume that distance or its proxy acts in an inverse manner, we can use the inverse power or inverse (negative) exponential. Sometimes the negative exponential has better mathematical properties

I just want to note that the typical model that we will discuss next week which is the basic spatial interaction model is in essence the same as this formulation except that we are going to deal with many origins and many destinations and we are going to add parameters that act as direct powers of variable. Then

$$T_{ij}\sim = T_{ij}\sim \frac{P_i^{\varphi} P_j^{\vartheta}}{d_{ij}^{\beta}} = P_i^{\varphi} P_j^{\vartheta} d_{ij}^{-\beta}$$

This will be the basis for the next two sessions which introduce the core ideas of spatial interaction

Scale and Agglomeration

Direct power laws also operate with respect to size. As an object gets bigger it changes in shape and this is called allometry. Positive allometry reflects an object getting more than proportionately bigger as it increase in size or as it grows. Negative allometry occurs when an object gets r smaller when it grows. Positive allometry is described by powers greater than unit whereas negative is described by powers 0 to unity.

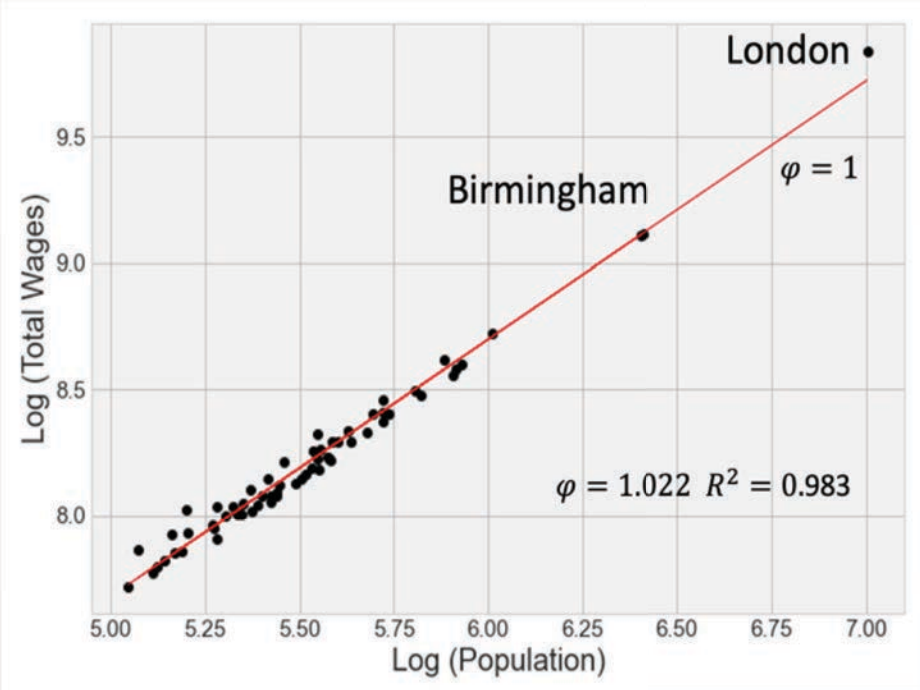
negative: $P(s) = Ks^{0.5}$ **positive** $P(s) = Ks^{2}$

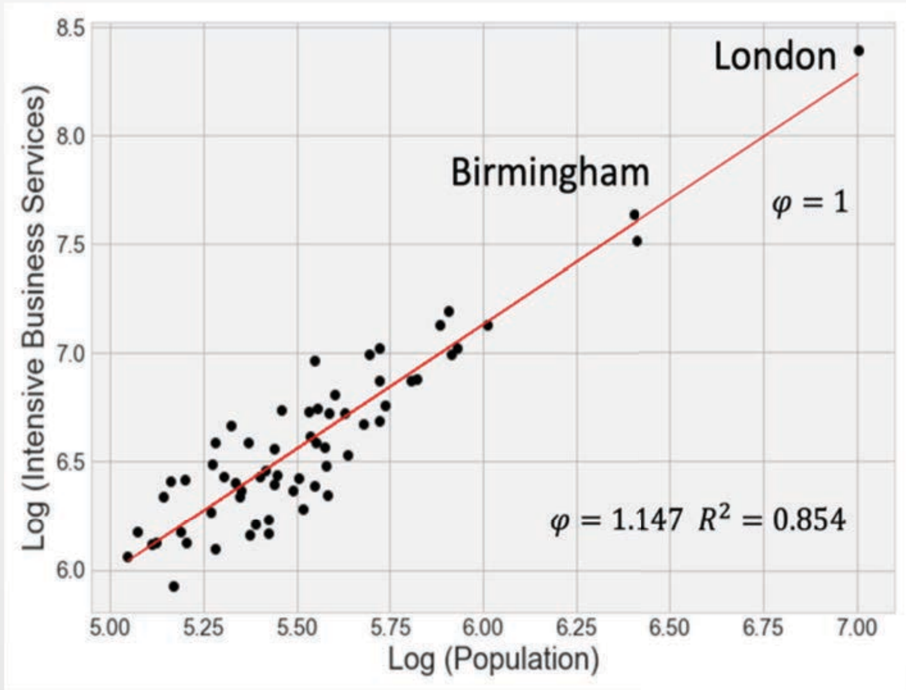
negative implies ***sub-linearity*** and positive ***super-linearity***

These effects are essentially economies of scale as developed in economics. They are sometimes called agglomeration economies and there can be positive economies of scale, and negative economies –often called diseconomies.

The effects have been widely observed with respect to the increasing size of cities and elements within. In particular Bettencourt and West have measured these effects are all over the globe in different cities. But the degree to which these relations depart from linearity is small and therefore controversial. Strong positive allometry is with $P(s) = Ks^{\beta}$ where $\varphi > 1.15$.

We can look at some of these effects for British Cities. The core 63 cities defined by the Centre for Cities has good data on employment, population and income We can graph these data as follows





Note that wages v population is linear despite what W and B say wile business services v population is super linear. Why?

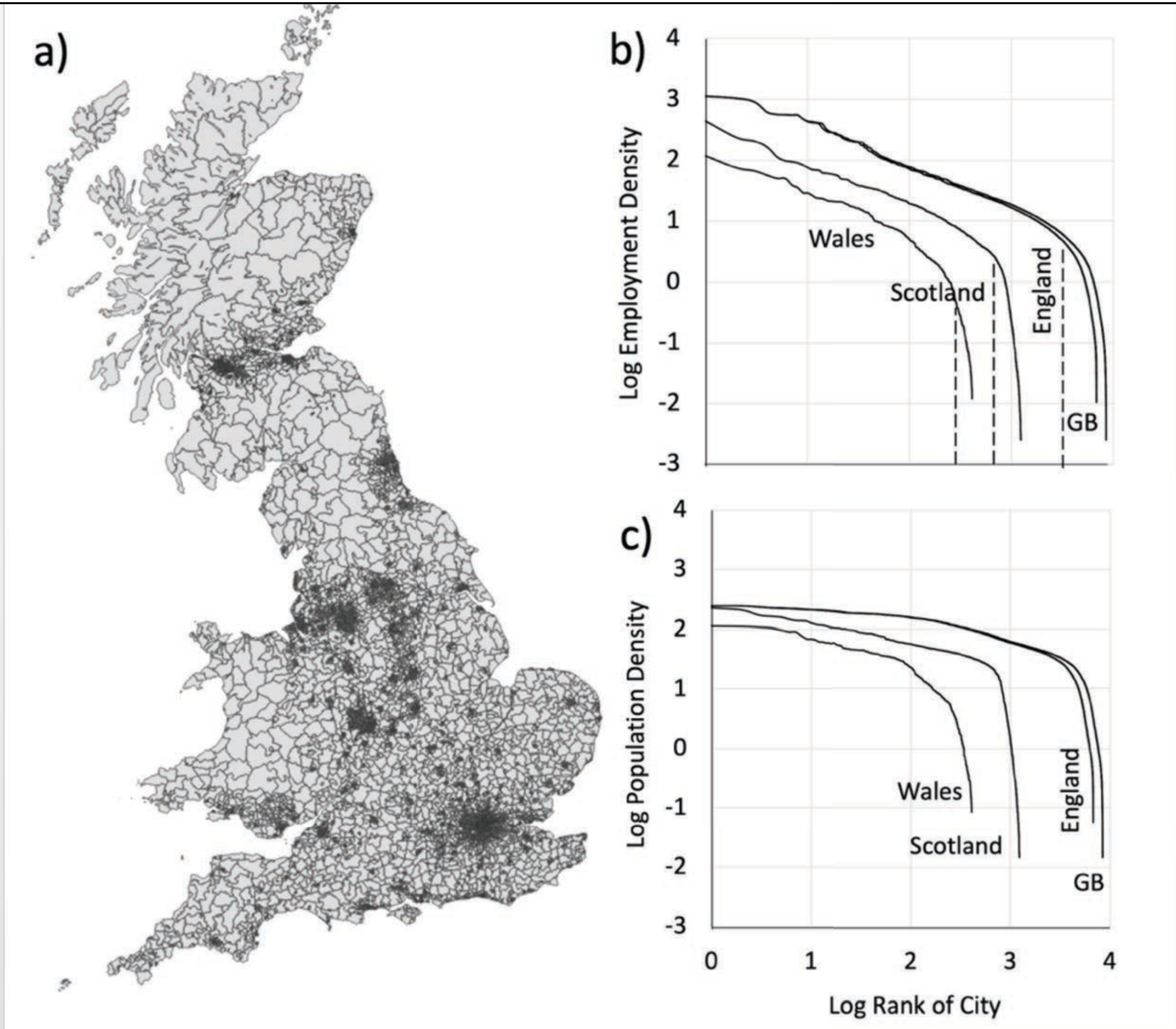
Rank-Size: City Size

A somewhat strange power law relates the size of cities to one another and it is essentially the frequency distribution of cities of different sizes. Cities sizes are not normally distributed (Gaussian) but are if anything lognormal but the biggest tend to be distributed according to an inverse power. In short, the number of cities increases as they get smaller and the logic for this is obvious. All cities cannot grow to be the biggest. There is competition for growth and as cities get bigger there are less of them. This power law can be described by ranking the cities from largest to smallest.

We now define the population P_r of a city at rank r and the rank-size law or rule can be defined as

$$P_r = Gr^{-\gamma}$$

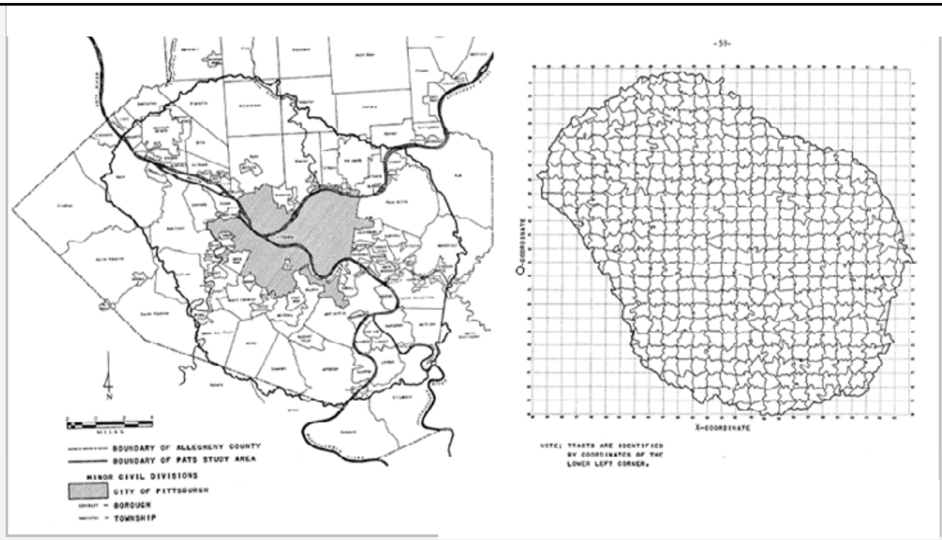
The parameters are the scaling constant G and γ and the most curious of all thinking about city size is that the γ is equal to 1 which means that $G = P_1$. In fact the power law only really operates in the heavy tail not the long tail but it could be argued that in fact the law should be for the positive side of an exponential distribution which approximates the inverse power. We show some of these distributions based on employment and population density for Great Britain.

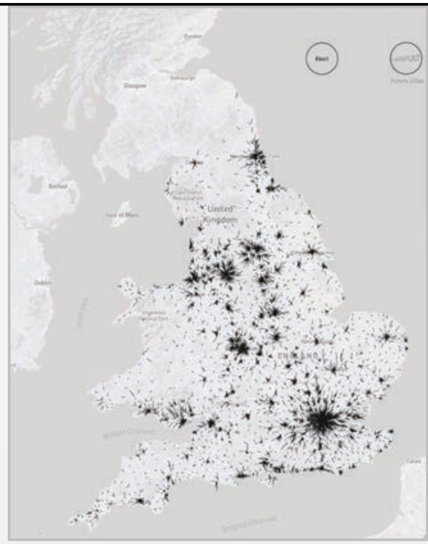

a)
b)
Log Employment Density
Wales
Scotland
England
GB
c)
Log Population Density
Wales
Scotland
England
GB
Log Rank of City

Generalising to Discrete Systems From Continuous

These ideas can be generalised in many ways. Power laws in network will be covered by Elsa but they work for all sorts of activities in cities and regions – even though we have only looked at population and employment.

Also we need to generalise all these equations to discrete systems as we shown for GB but here to finish let me simply show a typical gravitational model for a discrete zoning system using the same model to compute many flows between different locations





Vectors of flows positioned as average directions of all trips flowing into different destinations

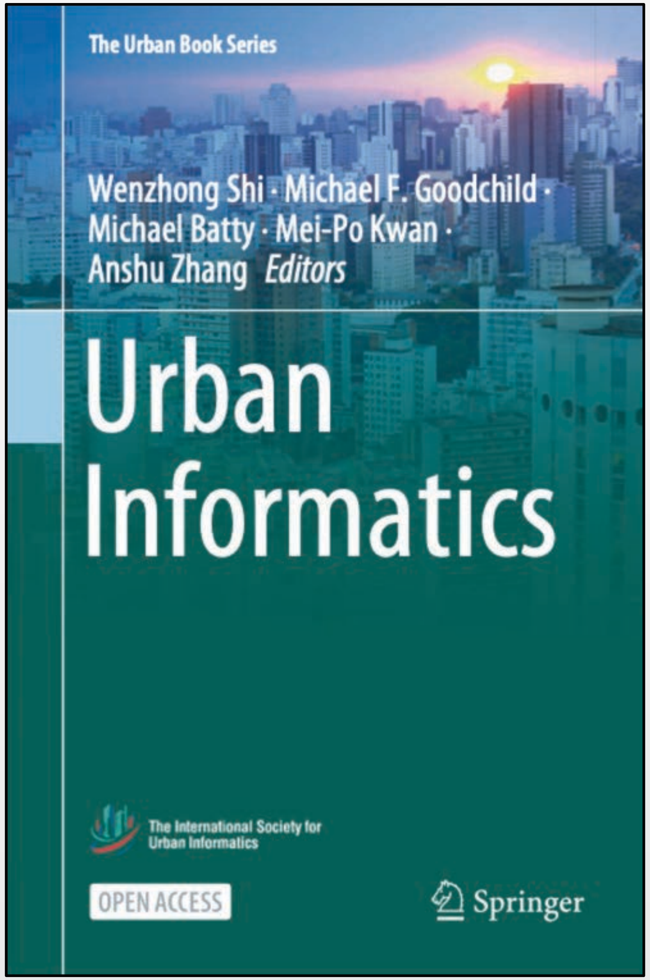
Simulation Models

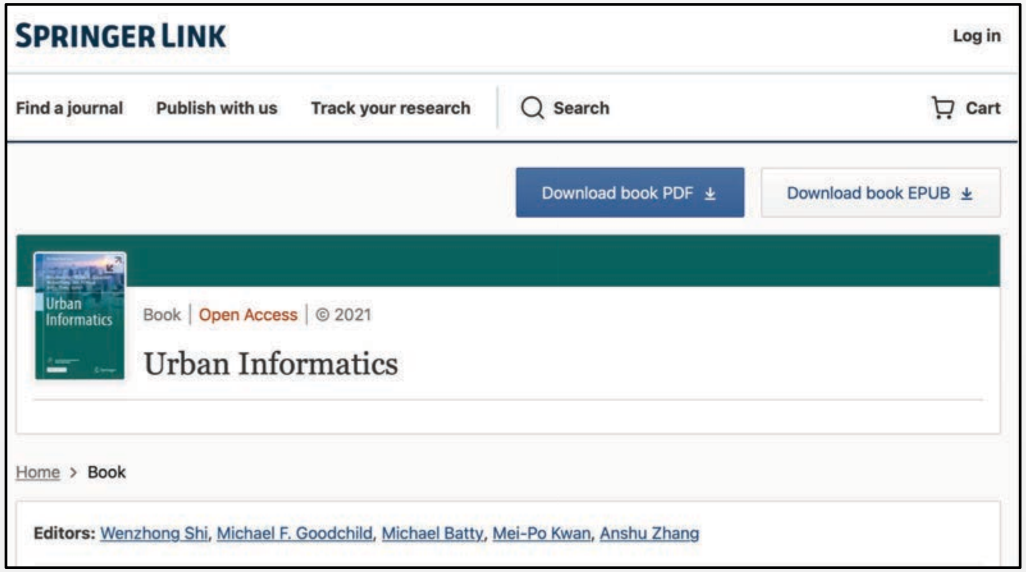
Reading for these lectures

The chapter in the Urban Informatics book by myself called "*Defining Urban Science*" is important reading and you can download it from

https://link.springer.com/chapter/10.1007/978-981-15-8983-6_3

The paper called "*Urban Scaling*" that is a draft of a chapter in a book on scaling can be downloaded from here:





Batty, M. (2021). Defining Urban Science. In: Shi, W., Goodchild, M.F., Batty, M., Kwan, MP., Zhang, A. (eds) **Urban Informatics**. The Urban Book Series. Springer, Singapore. https://doi.org/10.1007/978-981-15-8983-6_3