School of Geographical Sciences & Urban Planning

Lectures on Spatial Complexity
17^{th}-28^{th} October 2011

Lecture 3: 21^{st} October 2011

Simple Spatial Growth Models

The Origins of Scaling in Size Distributions

Michael Batty

m.batty@ucl.ac.uk
twitter @jmichaelbatty

http://www.complexcity.info/
http://www.spatialcomplexity.info/

Centre for Advanced Spatial Analysis UCL

Outline of Lecture 3

Scaling: Size Distributions: Frequencies, Power Laws and Ranks

Simple Models: Generators, Gibrat's Model of Proportionate Effect

City Size Distributions Observed: The US and the UK

The Idea of the Rank Clock

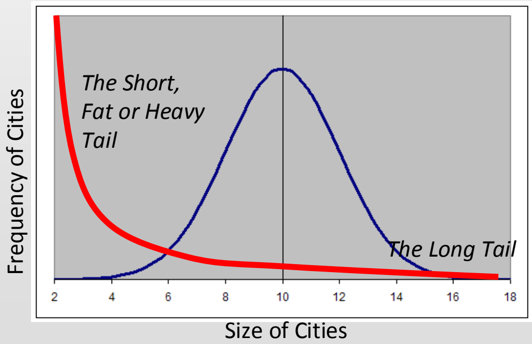
The US, UK, and World Urban Systems

Visualizing Size Distributions

Centre for Advanced Spatial Analysis UCL

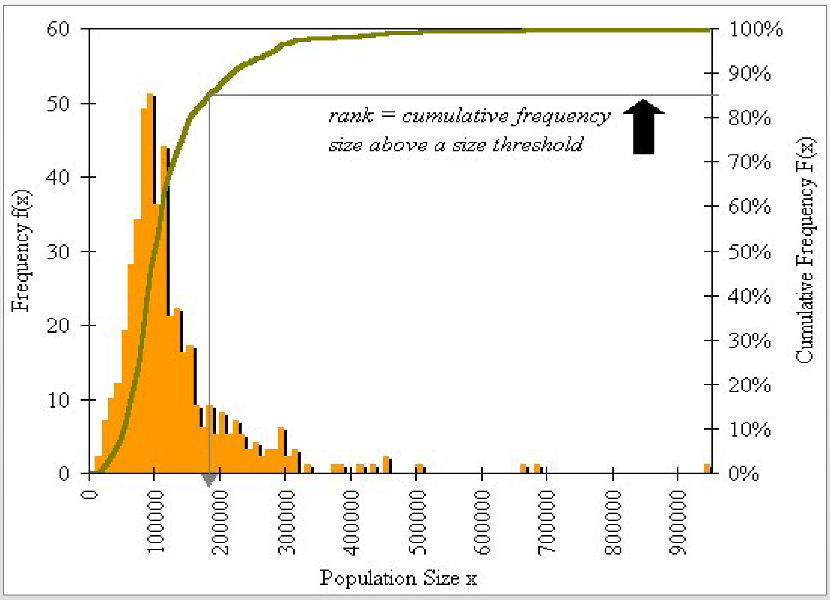
Scaling: Size Distributions: Frequencies, Power Laws and Ranks

Most distributions of objects in social systems where scarce resources and competition are intrinsic to their structure do not follow the usual Gaussian or normal distribution. They follow a power law



Centre for Advanced Spatial Analysis

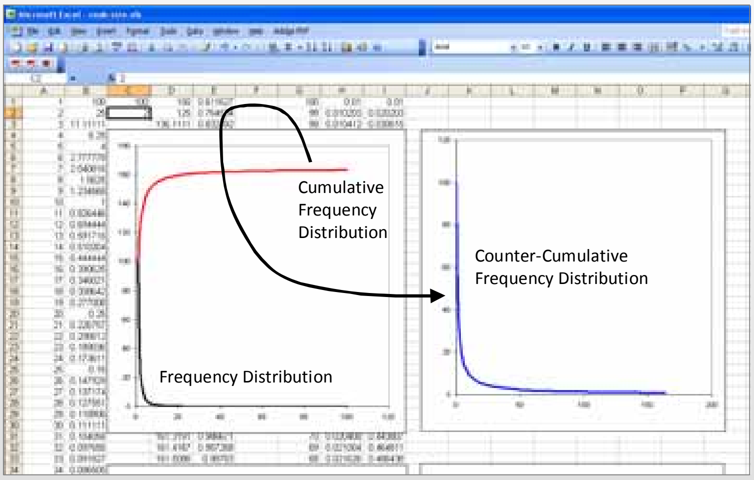
UCL



Centre for Advanced Spatial Analysis

UCL

We can transform the frequency to the rank size distribution



Of course such manipulations change the semantics of long, short, fat and heavy tails

The mathematics of these distributions is comparatively simple and we assume that the frequency follows a power law: First

$$f(P) = KP^{-\beta}$$

And then we have cumulative

$$F(P) = \int f(P)\, dP \sim KP^{-\beta+1}$$

And then we have the counter cumulative

$$r(P) = F(P') = \int_{P'}^{\infty} f(P)\, dP \sim KP^{-\beta+1}$$

We can simplify the counter cumulative to show that this too follows a power law: then

$$r(P) = KP^{-\beta+1} = KP^{1-\beta}$$

$$(KP^{1-\beta})^{\frac{1}{1-\beta}} = r(P)^{\frac{1}{1-\beta}}$$

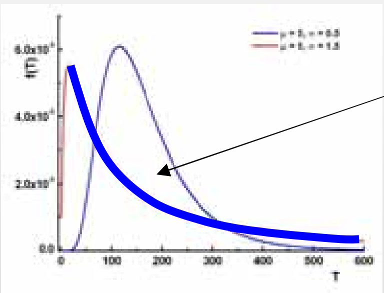
$$P(r) = Z\ r(P)^{\frac{1}{1-\beta}} = Z\ r^{\alpha}$$

This is called the rank-size distribution

$$if\ \ \beta = 2 \quad then \quad \alpha = -1 \quad and \quad P(r) = Z\ r^{-1}$$

more generally the rank size rule is written as
$P(r) = Z\ r^{-\alpha}$

In fact this is clearly an oversimplification – and usually the power law only holds in the upper tail. A vast array of evidence suggests that rank size distributions are more likely to be lognormal

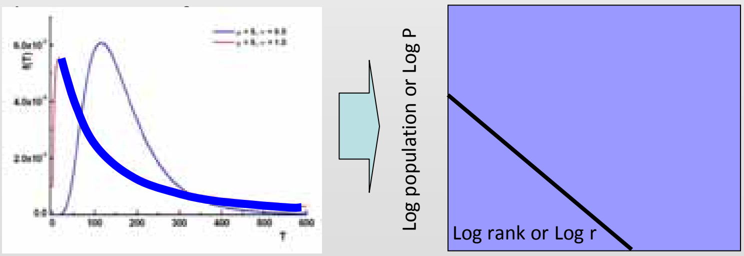


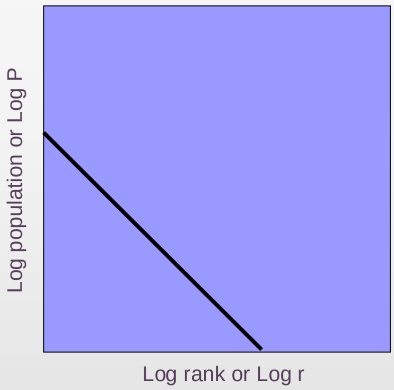
This one

But a lot of work on rank size is still fitting the power law to this upper tail

I could say an awful lot more about power laws, their limitations and their advantages – they are scaling – meaning that they imply all kinds of things about how systems grow and develop which are related to self-similarity and fractals.

In some senses, their apparent stability is a signature of self-organization and I will return to this. But the key issue here is how we fit them and it is easy to see that if we transform them to a log-log scale, then we get straight lines that make





The Strict Rank-Size Relation

$$P_r = P_1 \, r^{-1} = K \, r^{-1}$$

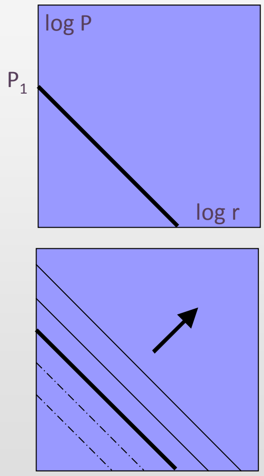
$$\log P_r = \log P_1 - \log r$$

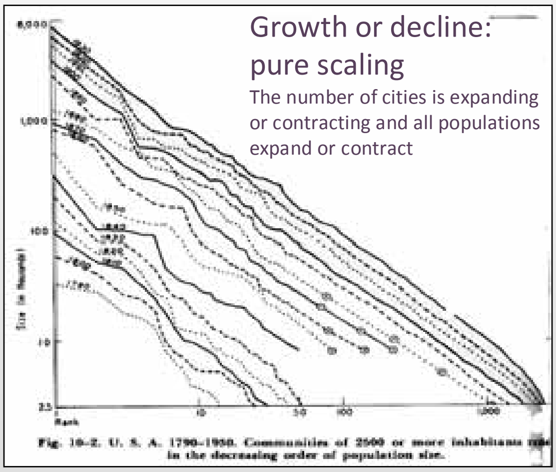
The first popular demonstration of this relation was by Zipf in papers published in the 1930s and 1940s.

The variable Rank-Size Relation which is the more usual is defined as

$$P_r = K \, r^{-\alpha} \qquad \log P_r = \log K_1 - \alpha \log r$$

There are clearly different types of systems such as



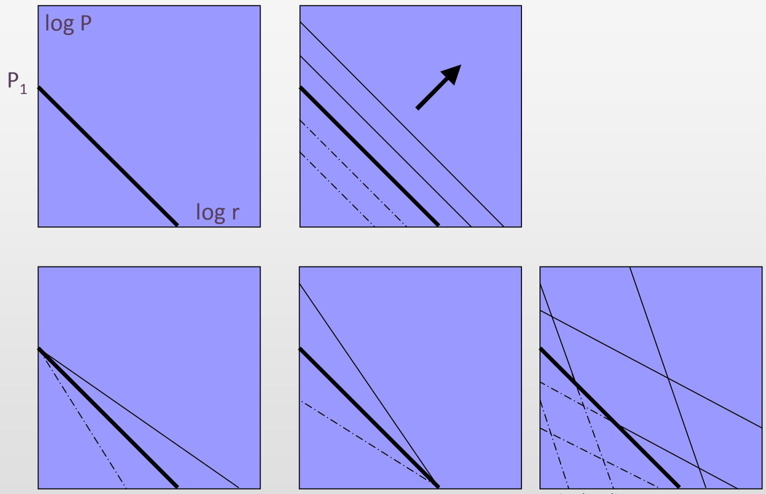


Growth or decline: pure scaling

The number of cities is expanding or contracting and all populations expand or contract

Fig. 10-2. U. S. A. 1790–1950. Communities of 2500 or more inhabitants in the decreasing order of population size.

There are clearly different types of systems such as



The number of cities is expanding or contracting and top populations are fixed.

The number of cities is fixed and all populations are expanding or contracting

mixed scaling:
Cities expanding or contracting, populations expanding or contracting

Simple Models: Generators, Gibrat's Model of Proportionate Effect

Now before we continue to look at the evidence for such scaling, then we need to consider how such ranking relations come about. Why is it that the number of large objects is small and the number of small objects is large?

On an intuitive level, then this is because of competition but why is the relationships so regular – why is it not more discontinuous or irregular with changes in slope?

To an extent, this is because of the repeated and regular action of competition – we will assume some of this is random and some of structured according to proportionate effects. But before we develop this model, let us look at a simple way of generating city sizes.

Now if we generate a new small city at each time period – and we let the cities grow exponentially say, then we only ever have one small city, one next big city and so on.

We don't have a rank size hierarchy and our frequency distribution is flat – we have one city of each size.

We thus need a mechanism that grows more small cities than big and one way of doing this is to assume that each time a cities spawn new ones.

A simple model is based on the idea that we start with a single city that then spawns another city at the next time period. So we start with 1 and then we have 2. At the next time period, the original city spawns a new city and the first new city spawns another, so we have 4 cities in all. Then with 4 cities, we spawn 4 more at the next time period making 8.

8 leads to 16, then to 32, and then 64 and so in the progression that can be modeled as

$$numberofcities = 2^t$$

Where t is time, so the sequence goes as 2^0=1 which is the first given city, 2^1=2 the first spawned city at time t=2, 2^2=4, 2^3=8, 2^4=16, 2^5=32 and so on up through time .

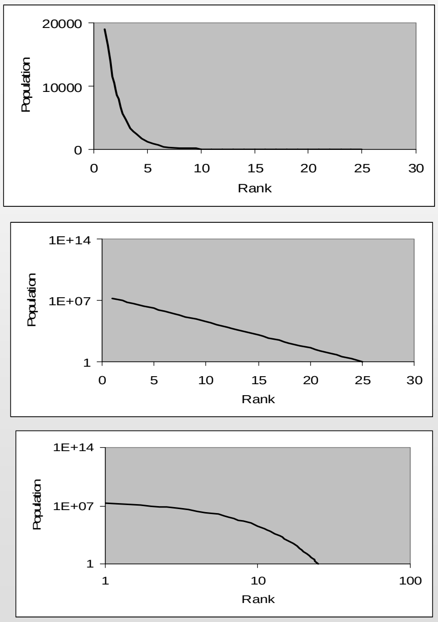
Now at each stage the city that already exists is not only spawning but it is also growing, and if it grows exponentially or nonlinearly, then it if we add all the cities together we will get classic exponential growth.

What we hope is that this process leads to a rank size rule or power law. In fact it doesn't as we will now show and thus was have to relax our model.

At each time, the cities grow as follows. In time 1, the start cities grows according to the rate and the new city is fixed as 1 unit. If the growth rate is say 1.5 (very large), then the city grows as 1.5*1. In the second time period, this start city grows as (1.5^2)*1, in the next as (1.5^3)*1 and so on.

By the time we get to the 10th time period, this start city is as large as (1.5^{10}) = 985, by 25th period, it is 33,537,598, which is fact is more than exponential growth. It is super or hyper exponential.

We have worked this out for this simple model and what we get is as follows: the frequency is exponential not power and thus this model gives massive weight to the biggest units. Moreover the model does not allow for cities to go up or down in size – and there are no deaths in the model.



We will show total population of the start city, then all cities, up to time 25 and then the population-rank relation, then the log pop-rank, then the log-pop-log-rank which shows that the relation is strictly an exponential not power law.

Clearly we can derived this from the process that we followed if we were to write the model in terms of its formula but we won't do this here.

Centre for Advanced Spatial Analysis

UCL

A much weaker but more acceptable model is not based on spawning cities which is rather an unrealistic hypothesis but assuming some sort of urban soup where cities rise and fall quite independently.

Imagine a landscape where at every point a city might grow (or decline if it is already established). The best way to think of this to begin with is if to assume every point is a tiny city of 1 unit size – a uniform distribution. Or we could assume that they are all randomly distributed in terms of size.

The uniform hypothesis is the best. If we assume that the growth rate of a city is proportionate to its size but that this rate is randomly chosen, then a city of P population at time t becomes P_{t+1} at time $t+1$ from $P_{t+1} = random(\lambda)P_t$

In fact we might write this more clearly as $P_{t+1} = (\lambda \pm \varepsilon_t)P_t$

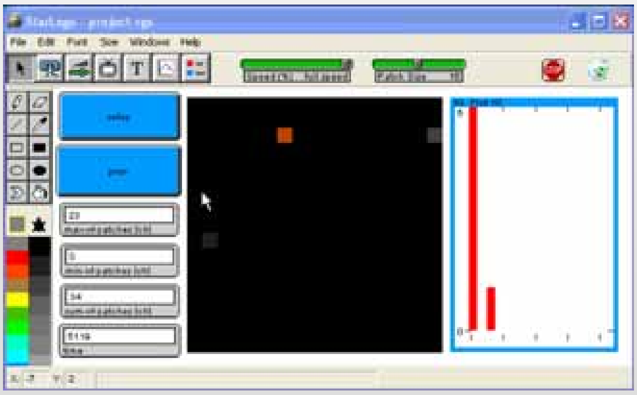
Centre for Advanced Spatial Analysis

UCL

Where λ is the growth rate and ε_t is the random effect. What happens to this process if we continue is that it is increasingly unlikely that a city gets big. The chances of the growth rate being large at every time period for any single city, decreases in time. In the same way, the chances of a city becoming very small are increasingly small but if we assume that cities cannot become negative – if they approach a population of zero, then this means that they either disappear or we do not keep them at a minimum level of 1 unit say.

In short, what this model is, is a 'random walk' with a barrier or threshold at a small number. The process is artificially constrained from the negative quadrant and thus competition is limited at the bottom of the pile. Negative values have no meaning as in the real world of cities

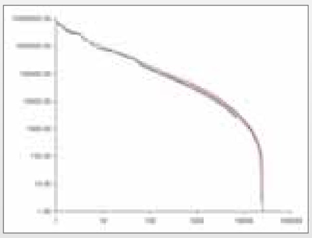
In fact I will show you a little movie of this process – I don't have the program handy but the movie basically makes the point that if you run this for long enough then what happens is that a rank size distribution – a power law of sorts emerges. Let me show you.

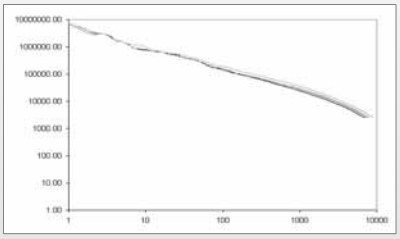


Centre for Advanced Spatial Analysis UCL

City Size Distributions Observed: The US and the UK

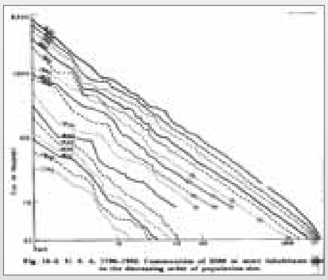
Ok let us move onto some empirical evidence – this has been a long digression and we need to take stock

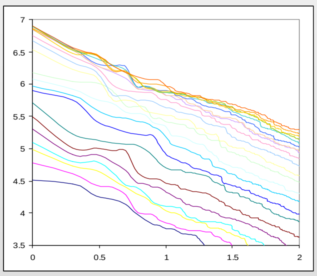




Parameter/Statistic	1970	1980	1990	2000
R Square	0.979	0.972	0.973	0.969
Intercept	16.790	16.891	17.090	17.360
Zipf-Exponent	**-0.986**	**-0.982**	**-0.995**	**-1.014**

I have begun to imply something about the dynamics. When we plot the log-log distribution for different periods of time, there is remarkable stability – quite remarkable. Here again is the US system from Zipf with a reworking to 2000





This remarkable stability, however is only skin-deep as we will demonstrate here

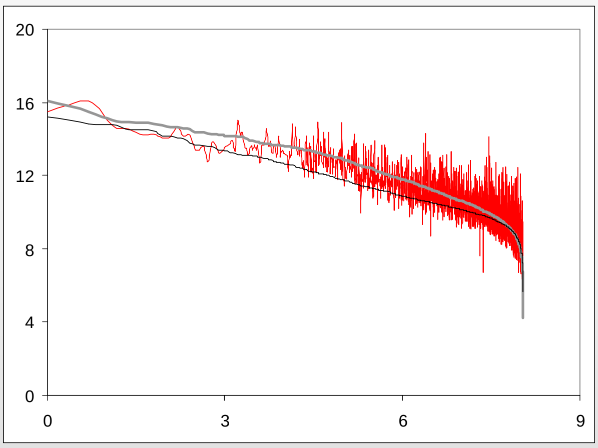
Year	r-squared	exponent
1790	0.975	0.876
1800	0.968	0.869
1810	0.989	0.909
1820	0.983	0.904
1830	0.990	0.899
1840	0.991	0.894
1850	0.989	0.917
1860	0.994	0.990
1870	0.992	0.978
1880	0.992	0.983
1890	0.992	0.951
1900	0.994	0.946
1910	0.991	0.912
1920	0.995	0.908
1930	0.995	0.903
1940	0.994	0.907
1950	0.990	0.900
1960	0.985	0.838
1970	0.980	0.808
1980	0.986	0.769
1990	0.987	0.744
2000	0.988	0.737

So if there is stability of this kind at the aggregate level and such volatility at the micro level, how do we understand it.

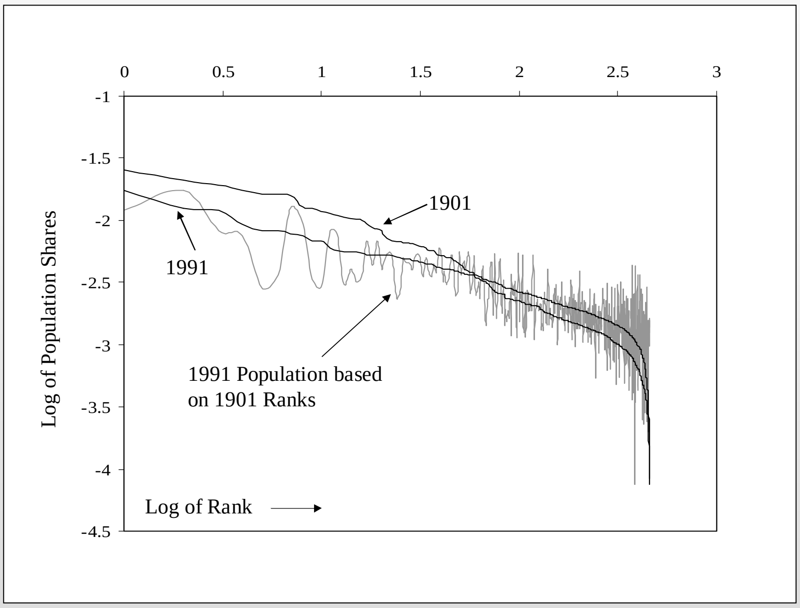
The first thing is to visualise it and the rest of this talk is about such visualisation

There are two devices we have developed: first the Rank Space – i.e. movement in terms of the position of cities on the original rank size or Zipf plot, and secondly a Rank Clock as we will show

But there is also first the rank shift that we have seen. This is merely plotting a distribution of rank size for a particular year using the ranks at a previous or later year. Here is the US from 1940 to 2000 with switches in rank order when we plot population of the 1901 cities with their 1991 ranks.



Now for the UK urban system from 1901 to 1991 but the other way around

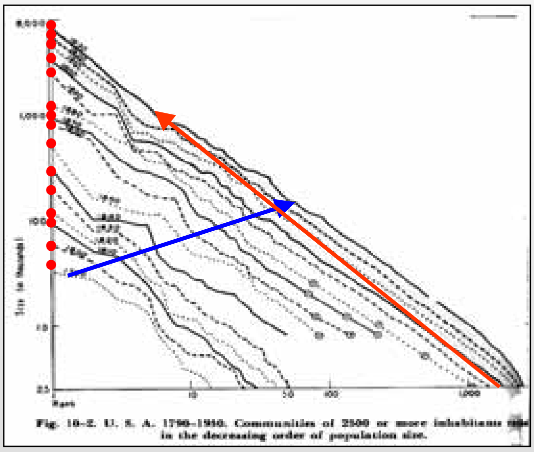


Centre for Advanced Spatial Analysis

UCL

The Idea of the Rank Clock (and Other Visual Mnemonics)

Log (Population)



Log (Rank)

- New York

Houston, TX

Richmond, VA

From George Kingsley Zipf (1949) ***Human Behavior and the Principle of Least Effort*** (Addison-Wesley, Cambridge, MA)

Centre for Advanced Spatial Analysis

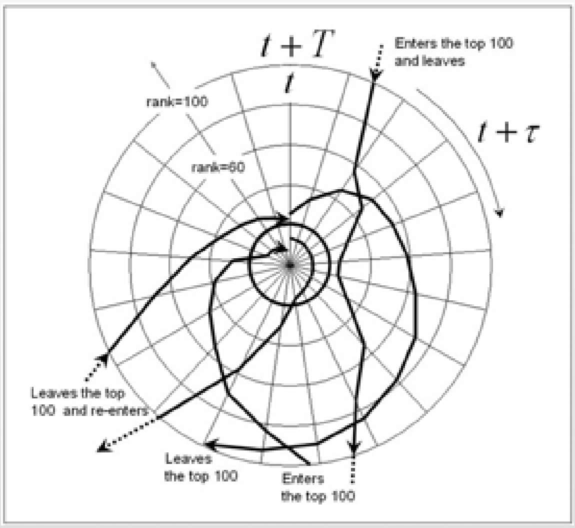
UCL

The idea of the rank clock is to discard population because rank is its equivalent and just plot changes in rank with respect to time

We can do this in rank space but it is not as effective as in polar coordinate space

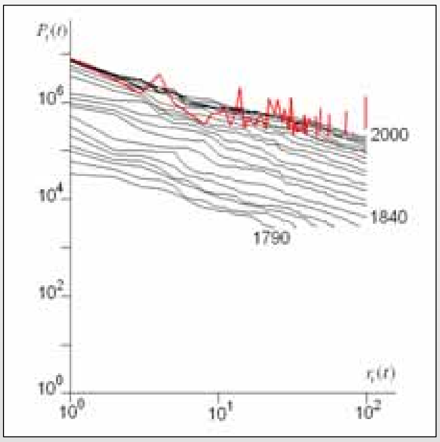
Here is an example of the possible trajectories of cities on the clock and we will stick with our US cities example for a bit longer, showing various possible plots and animations

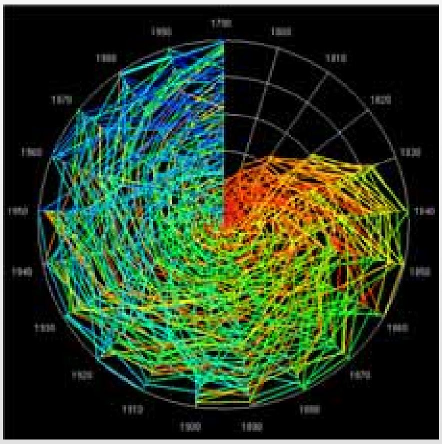
These clocks are useful for small numbers of cities but as we get more and more cities they become confused and we need to animate them in various ways as we will show below.



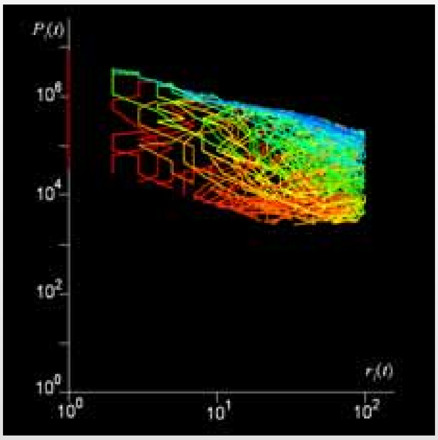
The Idea of a Rank Clock –rank is from number 1 at centre to 100 at edge and time goes in years in the usual clockwise direction

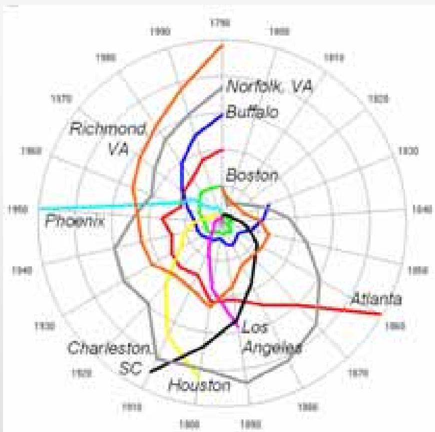
The US, UK, and World Urban Systems



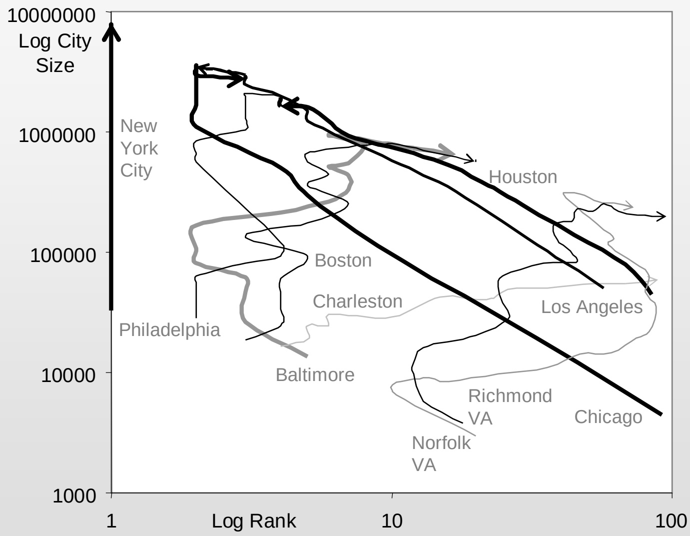


The ‘morphology’ of the clock should tell us something – i.e. the increase in cities, the volatility of ranks and so on.

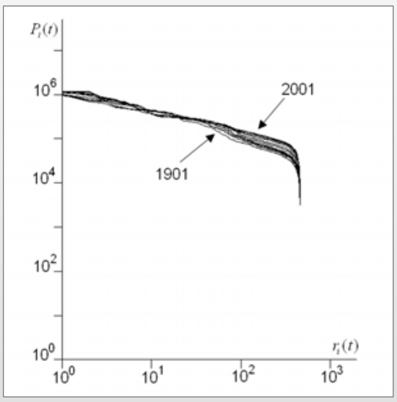


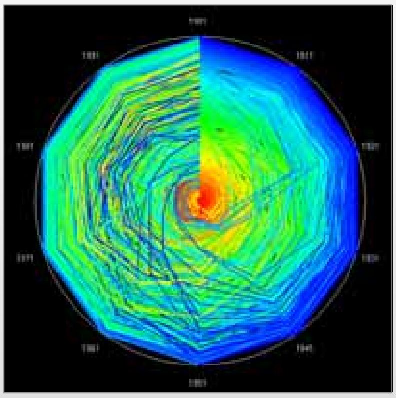


The rudimentary software for this in on our web site at
http://www.casa.ucl.ac.uk/software/rank.asp
And here is an animation of the US clock .. RankClockUSCities.exe



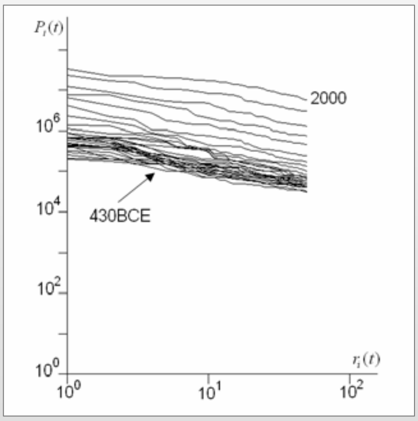
Another Example: The UK from 1901 to 2001

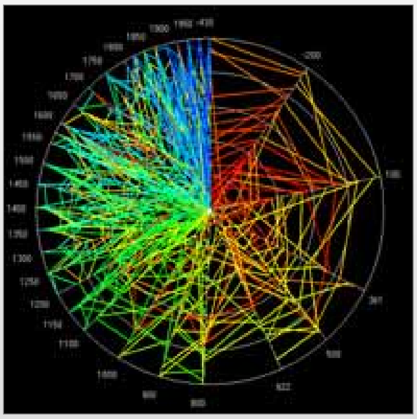




Very Long Term Dynamics: From 430 BCE

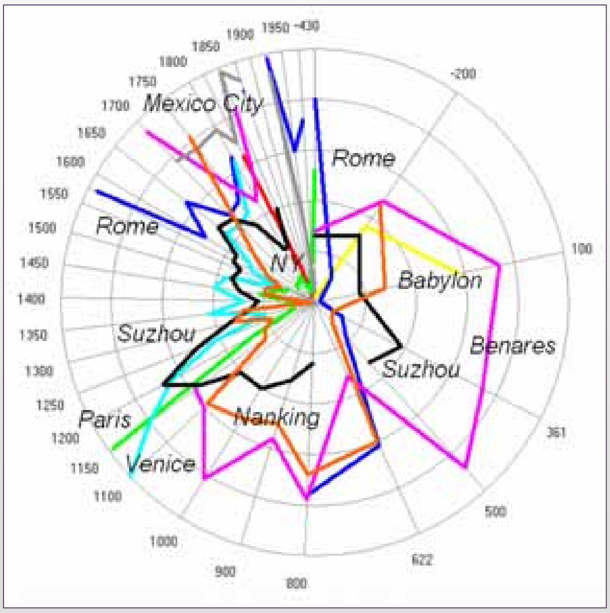
From the Chandler data set, for the top 50 cities we have





Centre for Advanced Spatial Analysis

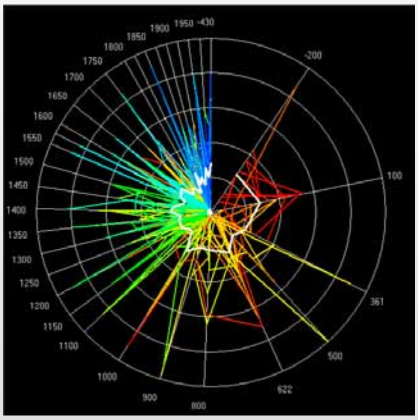
UCL

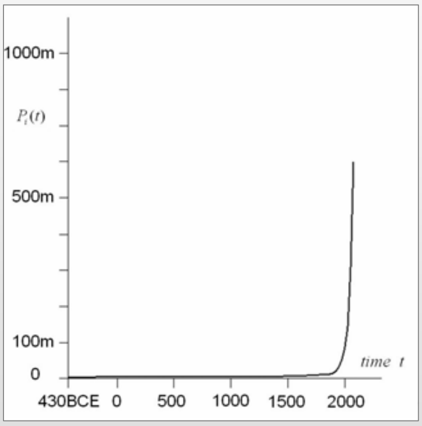


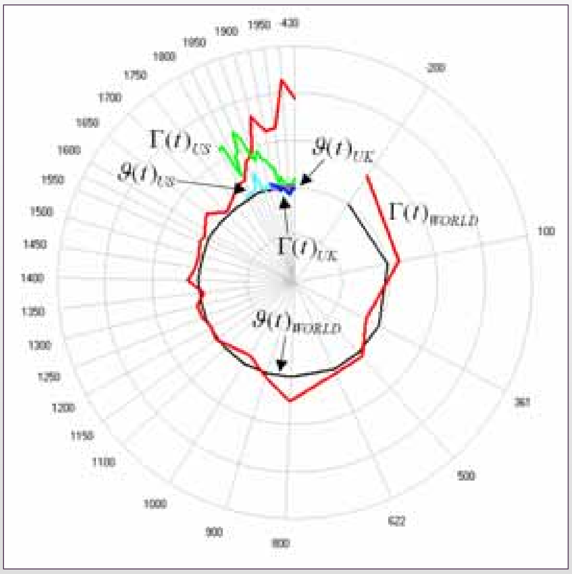
Centre for Advanced Spatial Analysis

UCL

Distances







Centre for Advanced Spatial Analysis UCL

Visualizing Size Distributions

We are currently producing web and desktop animations of these ideas and to conclude I will show and point to two areas of work that we are developing.

First we have a rank clock visualiser web page which is at

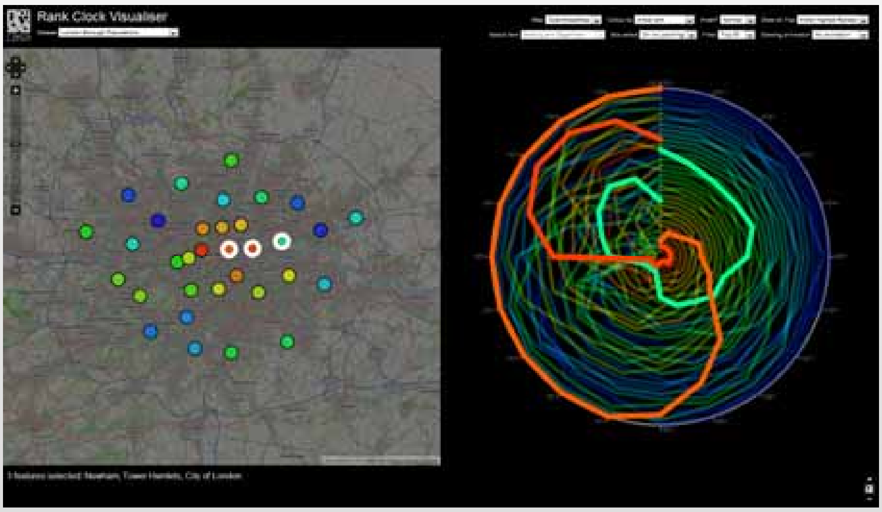
http://orca.casa.ucl.ac.uk/~ollie/rankclocks/

And we can link the dynamics of rank size to population distributions which are mapped spatially.

Centre for Advanced Spatial Analysis

UCL

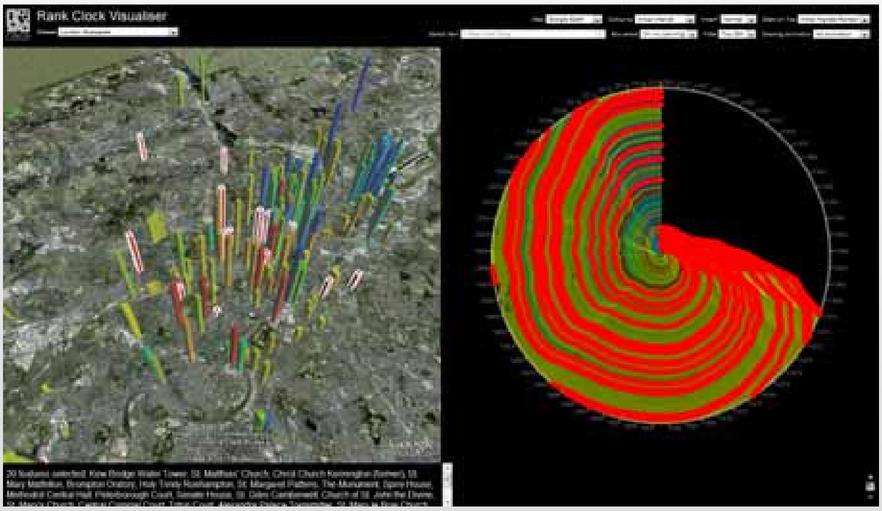
Demo – populations in 33 Greater London Boroughs from 1801



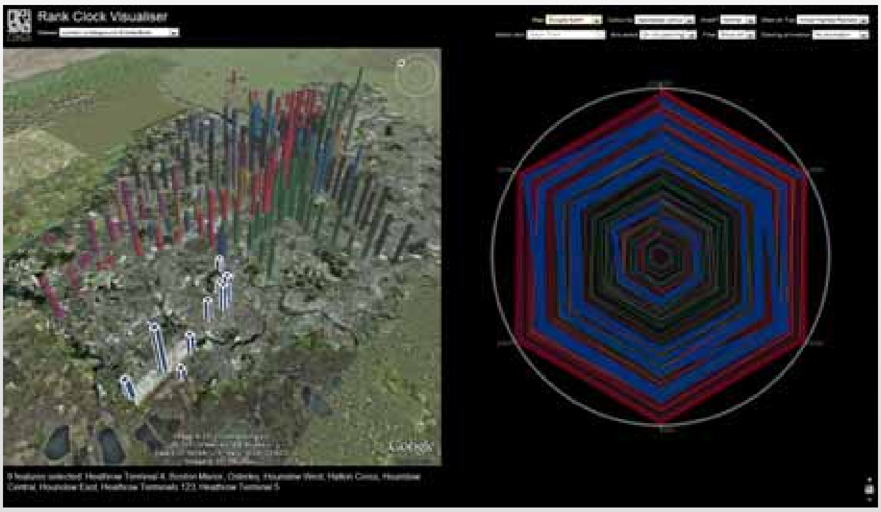
Centre for Advanced Spatial Analysis

UCL

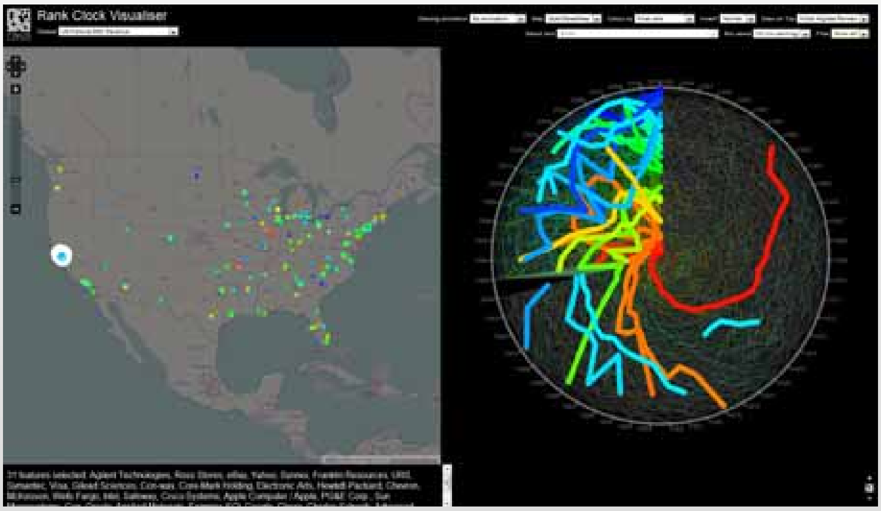
Demo – High buildings in London from 1950 to 2015



Demo – Tube Exit Volumes from 2003 to 2009



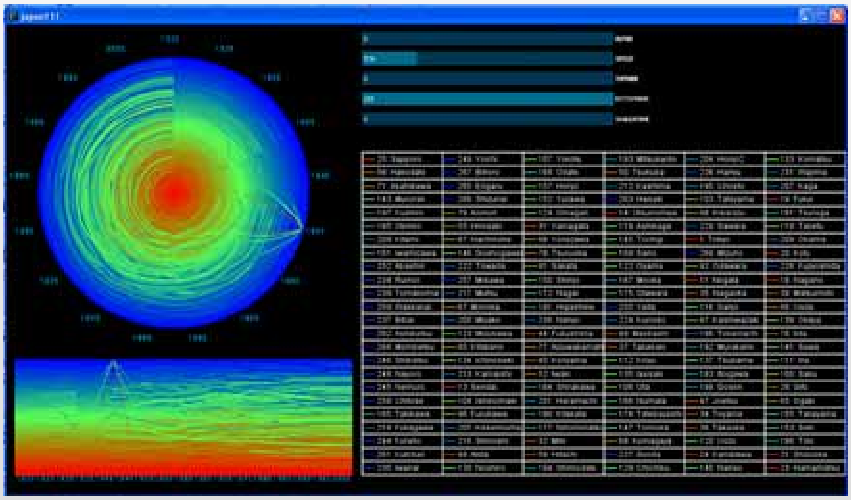
Demo – Fortune 500 from 1955 to 2010 for a sample of fast rising companies



Centre for Advanced Spatial Analysis

UCL

Here is the typical user interface in Processing



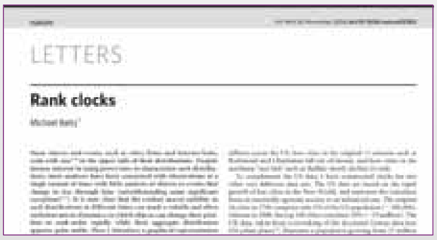
Centre for Advanced Spatial Analysis

UCL

Couple of Key references on this work but I will post more in the bibliography and refer you to material other than my own – if you click on these it takes you to my Science of Cities blog where the papers sit

http://www.complexcity.info/files/2011/06/batty-complexity-2010.pdf

http://www.complexcity.info/files/2011/06/batty-nature-2006.pdf


LETTERS

Rank clocks


RESEARCH ARTICLE

Visualizing Space–Time Dynamics in Scaling Systems

Centre for Advanced Spatial Analysis

UCL

The blog will have more and more references as the course continues

Questions

www.complexity.info

www.spatialcomplexity.info

Centre for Advanced Spatial Analysis

UCL