Environment and Planning A 2011, volume 43, pages 765 – 772

doi:10.1068/a43403

Commentary

When all the world's a city

Suddenly all the talk is about Chongqing, a city on the Yangtze which has reportedly reached a population of 32 million (figure 1). I say 'reportedly' because defining a city's boundaries is incredibly difficult and this astonishingly large figure relates to a very wide hinterland. Most world cities, however, have hinterlands of this size: 34 million in Greater Tokyo, more than the 21 million people in southeast England which is the true extent of Greater London, and 22 million in the New York Tri-State area which excludes Philadelphia. Chongqing is significant because it has reached this size seemingly without most people, even in China, knowing about it, and many never having heard about the very existence of the city. In fact, rather surprisingly, the authoritative *City Population* database (http://www.citypopulation.de/) records its population as only 6 million.

Figure 1. [In colour online, see http://dx.doi.org/10.1068/a43403] Chongqing: the world's latest megacity on the Yangtze.

All this confusion is occasioned by the massive migration to cities which began over 200 years ago and which appears to be still accelerating. This is nowhere clearer than in China where the number of cities with populations over 1 million is about forty but where it is predicted that by the year 2030 there will be 220. I have just returned from Guanzhou which now has a population of 10 million within the city boundary but arguably 30 million in its hinterland. When I first visited the city in 1986 it was a respectable 3 million, big then by Chinese standards. I had not been there for ten years and I barely recognised the place I had known quite well in the 1990s. A new central business district was virtually complete, having been constructed over a period of about four years, whereas back in 1986 there was no subway, there were barely any cars, just a sea of bicycles. Now the city boasts nearly 116 km of metrolines compared with the London tube which has 400 km but was built over 125 years not ten. This remarkable pace of urbanisation clearly cannot continue, for very soon the world's population will be completely urbanised and there will no longer be any possibility of moving to the cities. We will then all be living in some kind of city and the focus will be on migrations between cities which will comprise one globally interconnected urban system.

To get a sense of what is likely to happen in the next hundred years, look at the population of the top-fifty cities since the beginning of the classical age which dates from about 500 BC. This growth is superexponential. This population is shown in figure 2(a) from which it is quite easy to fit a growth curve extremely well to the past data (Chandler, 1987). If we superimpose the growth of the world's population on this, we see that since about 1850 this growth has been less rapid than that of the top-fifty cities, a fact that is dramatically illustrated when the ratio of the total population in the top-fifty cities to world population is plotted as it is in figure 2(b). Now at some point, if the growth of these two populations continues to follow these same trajectories, the top-fifty world cities will quite quickly outstrip the world's population. This is clearly nonsense. It cannot be the case. Something has to give. Let us consider what this might be.

The growth curves fitted to the two population groups are hyperbolic functions where the rate of change in population $dN(t)$ is proportional to some power α of the population itself $N(t)$, that is, $dN(t)/dt = N(t)^{\alpha}$. This model has the interesting nonlinear form $N(t) = N(0)/(1-\beta t)$, where a singularity appears when $t = 1/\beta$, that is, when the population goes to infinity. It was first proposed as a model for world population growth by von Foerster et al (1960) who predicted that the date when world population approached infinity (which they called 'doomsday') would be 13 November 2026. This is a date of mythical proportions in that the singularity

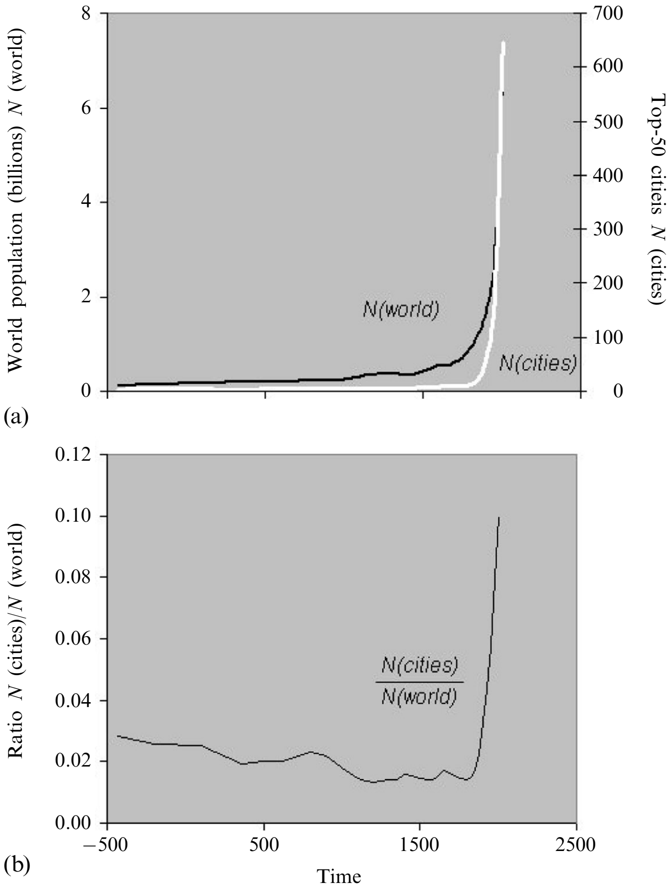

Figure 2. (a) Population growth (world and the top-50 cities); (b) ratio of the population of the top-50 cities to world population.

theorists—Vinge (1993) and Kurzweil (2005) amongst others—predict that at this date, or thereabouts, technologies will have enabled us to reach the point when, for all intents and purposes, machine intelligence will outstrip human intelligence, and man will become immortal; in population terms, the death rate will fall to zero. This model is fitted to both the world and top-fifty cities population using regression which is illustrated in figure 3(a). It is clear that, if we use data from the year 1400 from whence the linear trend becomes very pronounced, then for world population, the predicted 'doomsday' comes a little later than the mythical date of 2026: in fact on 4 May 2044. For the top-fifty cities population, however, the date is 9 February 2038, unerringly close to Kurzweil's (2005) own estimate of 2040. Brand (1999) refers to this singularity as "techno-rapture" and puts the date of this discontinuity at about 2035. The fact that doomsday is postponed for some eighteen years over and above the von Foerster limit is due to the use of new and more recent data, in this case, the augmented dataset originally compiled by Kremer (1993). The fact that this date is shifting forward is a simple recognition that such an event horizon is impossible. Or is it?

It is quite clear that these kinds of projections are nonsensical in real terms, but this is not the point. The point is that world population is continuing to increase at exponential rates and shows little sign of slackening—the demographic transition is barely to be seen in these data. The population of the top cities appears to be increasing faster than the world population and this is entirely consistent with the fact that the rate of urbanisation continues to accelerate. We are well past the mark

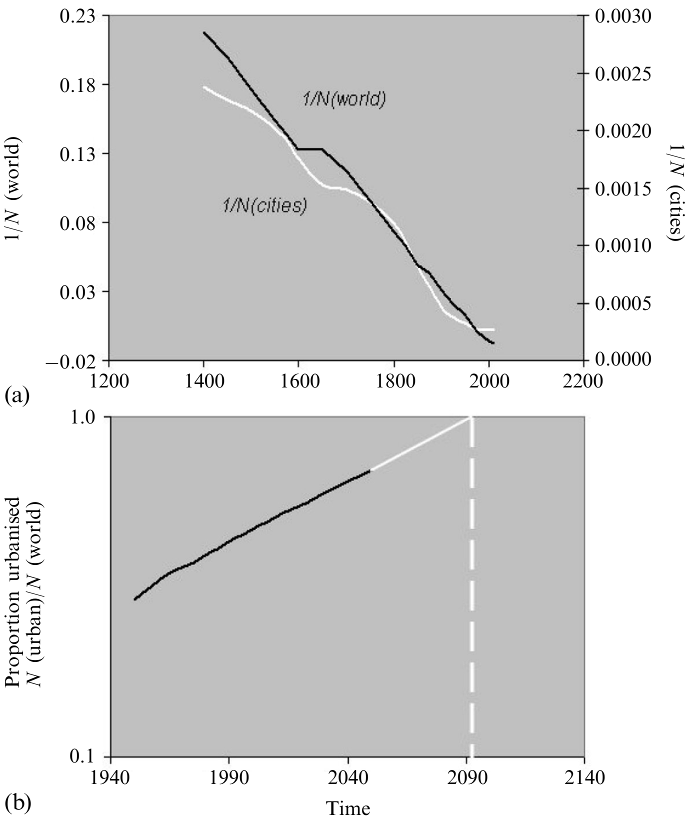

Figure 3. (a) Hyperbolic growth for world population and the top-50 cities; (b) the increasing rate of world urbanisation.

when more than 50% of the world's population was urbanised which took place in 2008, and independent UN predictions suggest the urbanised population will reach 70% by 2050 (http://esa.un.org/unup/). Figure 3(b) shows the remarkable stability of this urbanisation as predicted over one hundred years from 1950, fitting an exponential of the form $U(t) = K\exp(\lambda t)$, where $U(t)$ is the proportion of the population urbanised, with K and λ parameters of the growth equation. It is easy to show that, if this rate continues, the entire world will be urbanised by 2092. In short, all these trends with respect to cities and their urbanisation suggest that by the end of the century everyone will be living in cities.

But what kind of city? One city or many? When all the world's a city, the very concept of 'a city' begs redefinition. It is clear that no such thing as a completely connected physical entity called 'city' will exist unless we manage to shift continents which is unlikely. But a completely connected global entity where the connections are economic and social is already on the cards, in which it will be impossible to consider any individual city separately from its neighbours in economic space, indeed perhaps from any other city. To show how this might happen, we begin with the theory of superlinear scaling in cities developed by Bettencourt et al (2007) which suggests that, as cities grow, the resources $Y_i(t)$ they generate scale at more than linear rates with respect to their populations $N_i(t)$, where we now notate the location of any city by the index i. Thus $Y_i(t) = \alpha N_i(t)^{\psi}$ where the power $\psi > 1$. The logic of this relates to the fact that, as the population increases, the number of people generating resources that apply to the entire population increases in constant proportion to the population, but the effect of their actions increases the wealth of more than these individuals. In this sense, technologies that generate wealth are all pervasive or at least apply to (many) more people than just their inventors. If everyone generated resources that applied to the entire population, there would be $N_i(t)^2$ such effects; in other words, everyone would influence everyone else. This is unlikely, for it is clear that the number influenced by these wealth-generating effects would be less than the total population; hence the logic of $N_i(t)^{\psi}$. In fact, only a constant proportion α would generate these kinds of interaction effects, thus leading to the resources equation which, when fitted to the largest US cities, gives $\psi = 1.12$. As Bettencourt et al (2007) show, there are many technologies and creative pursuits that scale in this manner and this is consistent with the argument that, as populations grow, this induces endogenous technological development (Kremer, 1993).

The resources in any city are also considered as a function θ of maintaining the existing population and, together with the addition of resources ϕ to support an additional individual, we can express this as $Y_i(t) = \theta N_i(t) + \phi[dN_i(t)/dt]$. Now, if we equate this with the wealth equation giving $dN_i(t)/dt = (\alpha/\phi)N_i(t)^{\psi} - (\theta/\phi)N_i(t)$ it is easy to show that this differential is similar in structure to the equation we used to model the world's population above. In other words, if an individual city's wealth increases superlinearly with population, then this implies that the growth dynamic of population follows a hyperbolic curve which leads to exactly the same sorts of singularity that were demonstrated above. Bettencourt et al (2007) resolve this issue for individual cities by suggesting that the technological clock is reset every so often as cities head towards the point of singularity because other limits come into play. Our focus here, however, is not on the resetting of the clock which has a clear enough logic but on how cities begin to interact in forming the kind of global entity that constitutes the state when all the world's a city. For this, we now need to consider how cities interact and connect with one another as they grow.

There are two very well-tracked and rather obvious assumptions relating to urban growth involving centralising and decentralising forces which we can characterise as attraction and diffusion. Attraction relates to the fact that cities suck up population from their surrounding hinterlands, as in the case of Chongqing, but also pull in people globally as they extend their reach. This reach is in fact that diffusion effect which, I argue, grows ever greater as the city gets bigger, eventually extending either directly, or more likely indirectly, to link all the world's population. To sketch out what this means, we need to consider the entire system of cities at discrete locations $i, j, , \ldots$.

Next, we add a migration component m_{ij} to the differential which will reflect the physical distance d_{ij} to other cities as a deterrent to such flows taking place, but which is countered by an attractor between the two places in question that is some function of their respective populations. We can specify this in the form used in migration modelling since Ravenstein (1885), and articulated, for example, as $m_{ij} = N_i N_j / d_{ij}^2$, noting again the appearance of powers of population in the growth. This diffusive force is not a repellent for this reflects the fact that, as cities get larger, they exert an economic (and social influence) on other cities that materialise in terms of their economic connectivity. Undoubtedly, there might be additional growth effects that occur when cities begin to connect up, but at present all we need to worry about is the point at which cities connect to form part of the endemic global cluster that I suggest occurs when all the world's a city. It is probable that, when cities connect up to one another, some threshold Λ, involving the potential for trade, is reached and we might specify this as a point at which they join the global cluster $G(t)$. We can thus hypothesise that, if two populations reach this threshold; that is, if two populations $N_i(t)$ and $N_j(t) > \Lambda$; then $G(t) = G(t-1) + N_i(t) + N_j(t)$. This can be generalised to varying thresholds which might be temporally grounded in terms of relative sizes and extended, of course, to more than a single pair of cities joining the cluster at a given point in time.

It is therefore quite straightforward to show that, as there are no clear deterrent effects in the augmented model, then $G(t)$ simply gets bigger with time while the number of cities that are not part of the global cluster continues to fall. All this is predicated on the basis that we have a fixed number of cities or city locations in the world: which is not as simple an assumption as one might imagine because cities can grow anywhere and we have not yet embraced the third dimension to any real extent. The singularity may be as much pushed back by the expansion of cities into this third dimension as it might be invalidated by the movement of populations on a mass scale off the planet. Such speculations are fanciful, the stuff of science fiction, but they are also clearly imaginable: Terminus City in Asimov's *Foundation* novels and the Death Star in the *Star Wars* movies capture up populations that have expanded into space very different from the usual limits on city form that we suppose and impose.

There are a thousand qualifications we need to make to this argument to inject more realism into it. The uniformity which is assumed in this dynamic is not realistic, for we know that cities also fall in size, and the model portrayed here does not allow this to happen. Moreover, all the time, cities move up and down the city-size distribution. It is quite likely that the superlinearity that is at the basis of the growth model varies between different places and different historical periods, reflecting intrinsic technological developments of particular eras. We could demonstrate all this by setting the parameters in the above equations as averages and then introducing variations around these means to take account of the heterogeneity in the world system. There are no simple equations which produce clean or closed solutions, but setting up such simulations would lead to continued growth of the global cluster while variations in individual city growth would also occur.

What we are suggesting is that, as cities grow and as their mutual attraction increases through migration and other network effects, they begin to coalesce in functional terms. However, this coalescence cannot happen from the top down, it must surely happen from the bottom up. In the world of cities, what appears to happen is that, as cities grow, their influence through migration on their surrounding populations reaches a point at which the city traps small cities into its hinterland, thus forming local and regional clusters of more tightly connected elements. Thus in the primitive stages of the evolution of a world system of cities, local structure appears. But, as some cities grow faster than others, eventually the threshold is passed when really large cities within their clusters join up. If one imagines these connections as a large network or graph, there comes a point at which a giant component emerges when the connectivity of cities is such that it is possible to consider every city to be connected to every other city, but with most of these connections being indirect, down through the size hierarchy. When this happens, a phase transition takes place which is revealed by the fact that the 'economic' path length in the graph goes from 'infinity' which is the state where some cities are not connected to the core directly or indirectly, to a state where all cities become connected. This is a familiar problem in percolation, but it maps rather well onto a simple analogue of how a global system emerges (Kali, 2003).

To close this argument, we need to demonstrate that there is such a thing as a global cluster which results when individual cities merge in this way. In one sense, this was anticipated at the very outset of this commentary when the set of top-fifty cities was introduced and it was noted that the population of this set was growing faster than the world population. In fact, what we need to show, if our hypothesis holds up at all, is that this set of top cities is actually growing in number and for this we need many more cities. To conclusively demonstrate this, we need *all* cities for we can never show this just for a fraction of cities at the top of the size distribution. How do we proceed? A very brave and somewhat tortuous hypothesis is to assume that there is a combined migration and diffusion effect which we can capture in the standard migration equation. I will argue that there is a basic threshold θ which, once reached, connects any centre of population i to any other j; that is, if $m_{ij}(t) \geqslant \theta$, then a local cluster $\mathrm{G}_{ij}(t+1) = \mathrm{N}_i(t) + \mathrm{N}_j(t)$ is formed. Imagine this process occurring when all cities are growing on average, which is quite consistent with the growth of world and top city populations shown earlier. What happens is that successive clusters merge, that is, $G_{ijkl}(t+1) = G_{ij}(t) + G_{kl}(t)$ when $\mathrm{m}_{ij}(t)$, $\mathrm{m}_{ik}(t)$, $\mathrm{m}_{il}(t)$, $\mathrm{m}_{jk}(t)$, $\mathrm{m}_{jl}(t)$, and $\mathrm{m}_{kl}(t) \geqslant \theta$, with an appropriate generalisation to all populations and all clusters. What we have to do is find this threshold θ and then we can easily look at the top-fifty cities over the last 2500 years and work out almost by inspection whether or not the giant cluster is emerging and growing.

Two points of caution. We are assuming that physical distance is critical and we know that this may not be a strong hypothesis in a world of information flows. Nor do we know how stable this threshold might be. However, the only way we can proceed is to establish some sort of threshold which appears plausible and then test to see how many of our top-fifty world cities are interconnected in this way, and how fast these interconnections are growing. This is effectively suggesting that the emerging global entity $\mathrm{G}(t)$ is that set of cities (and their clusters) whose migration flows between one another are greater than θ. One way of doing this is to look at a sample of cities at another level and 'guess' the extent to which they are connected; that is, to figure out if the intercity migrations are greater than θ. We can do this quite easily for US cities especially as several large cities in the US have developed over the last 200 years with the opening up of the American West and South. Examining the top cities in the US such as New York, Chicago, Los Angeles, and Houston provides us with some sense

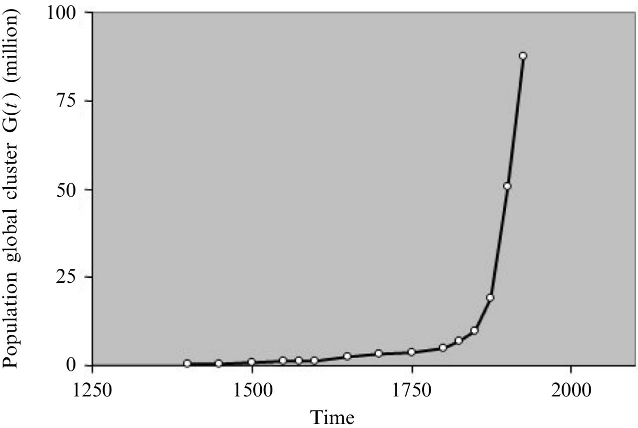

Figure 4. The population of the emergent global cluster from the top-fifty cities.

about the point at which these cities have connected to the emergent US cluster. On this basis I have computed the migration flows between them and consider a plausible threshold is $\theta = 50\,000$ where the migration equation above is used with raw populations and distances between these cities in kilometres. With this value we connect New York with Chicago, Los Angeles, and Houston in 1860, 1890, and 1920, Chicago with Los Angeles and with Houston in 1900 and 1930, and Los Angeles with Houston in 1940.

Now we can apply this to the top-fifty world cities from 1400. The cities of the Middle Kingdom first connect up with the eastern Roman Empire from the time of Marco Polo and then the European cluster begins to emerge in the late Renaissance period. Dramatic growth begins around 1800, but it is not until the mid-19th century that the Americas enter the cluster. In the set of the top-fifty cities I estimate that all of these are connected by 1925 and then when these populations are plotted, the globalisation profile shown in figure 4 is produced. The growth curve that can be fitted to this is also superexponential and suggests that the population of the cluster has already (in 2010) outstripped the world's population, with global urbanisation running well ahead of the curve. All this suggests that we may all live in cities by the end of the century, but there is sneaking suspicion that some may still not be connected to the giant cluster. The implication is that, in time, all cities will be connected, but there may remain stubborn pockets of resistance in failed states, where there are really poor, undeveloped parts of the world. A definitive demonstration must await more considered simulations, but this commentary suggests that this is possible. Of course, this is only the merest sketch of what we might do, but we urgently need to think out of the box with respect to what the world will be like when it is entirely urbanised, when all the world's a city. I would hope that this might be one good way to start.

Michael Batty
Centre for Advanced Spatial Analysis, University College London

References

Bettencourt L M A, Lobo J, Helbing D, Kuhnert C, West G B, 2007, "Growth, innovation, scaling, and the pace of life in cities" *Proceedings of the National Academy of Sciences* **104** 7301 – 7306

Brand S, 1999 *The Clock of the Long Now: Time and Responsibility* (Basic Books, New York)

Chandler T, 1987 *Four Thousand Years of Urban Growth: An Historical Census* (St David's University Press, Lampeter, Pembrokeshire)

Kali R, 2003, "The city as a giant component: a random graph approach to Zipf's law" *Applied Economics Letters* **10** 717 – 720

Kremer M, 1993, "Population growth and technological change: one million B.C. to 1990" *The Quarterly Journal of Economics* **108** 681 – 716
Kurzweil R, 2005 *The Singularity is Near* (Viking, New York)
Ravenstein E, 1885, "The laws of migration" *Journal of the Statistical Society* **46** 167 – 235
Vinge V, 1993, "Technological singularity", a paper presented at the VISION-21 Symposium, NASA Lewis Research Center and the Ohio Aerospace Institute, 30 – 31 March, http://www-rohan.sdsu.edu/faculty/vinge/misc/WER2.html
von Foerster H, Mora P M, Amiot L W, 1960, "Doomsday: Friday, November 13, AD 2026" *Science* **132** 1291 – 1295