

The shape of future cities: Three speculations

Michael Batty

Centre for Advanced Spatial Analysis (CASA), University College London (UCL), UK

Transactions in Urban Data, Science,
and Technology
1–6

© The Author(s) 2022



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/27541231221113945

journals.sagepub.com/home/tus



Abstract

World population has grown dramatically since the Industrial Revolution began 250 years ago. Cities are key elements in this growth, but by 2100, we will all be living in cities of one size or another. Here we speculate what this world will look like. First, the Industrial Revolution represents a clean break from a past composed of hardly any cities to one which is completely dominated by cities. Second, cities will continue to change qualitatively as they get larger but size limits will emerge. Third, cities will no longer be classified by their size but by their internal dynamics.

Keywords

Urban, technology, the Fourth Industrial Revolution, future, cities

A little after the last Ice Age ended some 12,000 years ago, man gradually moved from being nomadic hunter-gatherers to engage in more settled agricultural pursuits. It took another 6000 years or so for the first cities to emerge, although even by the beginning of the Common Era, only about 5% of the world's population lived in anything we might consider to be a city. World population had bumped along at no more than a gradually increasing few millions throughout most of human history and only after the Renaissance in Europe began during the 15th to 17th centuries was there any sign that exponential growth in the population might suddenly take off (Wiki, 2022). But take off it did, largely due to the fact that the Industrial Revolution spawned many cities which became jewels in the industrial crown, where new technologies were both invented and disseminated. Trying to explain the origins of the Industrial Revolution and why it did not begin long before from the chronologies of previous Chinese, Greek and Roman empires, for example, is not a quest that is easy to unravel. But there is little doubt that by the middle of the 18th century, the conditions in Western Europe were right for innovations that would lead to a succession of new technologies that, in turn, would raise prosperity for many above bare subsistence levels for the first time in human history. In hindsight, the onset of this revolution increasingly appears to mark a clean break with the past, as we will elaborate in more detail below.

The first Industrial Revolution, based on the invention of steam power, enabled populations to break the limits on distance that had dominated all previous societies. For the first time, a growing urban population could live at greater distances from their work but still remain part of the city, enjoying more space and less

Corresponding author:

Michael Batty, Centre for Advanced Spatial Analysis (CASA), University College London (UCL), 90 Tottenham Court Road, London, W1T 4TJ, UK.

Email: m.batty@ucl.ac.uk; t@jmichaelbatty

congestion. Cities could grow beyond their previous technological limits that kept them small, which for most was much less than the upper limit of about one million in population. The first Industrial Revolution introduced many new kinds of manufactured machines and by the end of the 19th century, a second revolution fusing mechanics with electricity led to the automobile as well as various devices that enabled populations to communicate easily at a distance such as the telephone and the radio. The second revolution also broke down the barrier that to communicate you had to physically move, and it quickly morphed into a third, the digital revolution. This was marked by the invention of the computer while a fourth revolution now underway is where information and digital technologies are penetrating all kinds of machines as well as ourselves, establishing a digital skin to the planet as well as enabling us to communicate with anyone in any place and at any time.

This labelling of successive Industrial Revolutions as chronologies is due to Klaus Schwab (2016), who also makes the point that each revolution does not cancel out or replace its predecessor but builds cumulatively on all those so far. In short, each revolution weaves its way into those that already exist, renewing previous innovations from within as well as from without. This is particularly the case with the third and fourth revolutions, which are complementing flows of energy with flows of information. Moreover, other futurists writing about these revolutions use similar terms. Toffler's (1980) *Third Wave* is consistent with Schwab's third and fourth revolutions, while the term 'post-industrial', associated with Daniel Bell (1974) amongst others, is also applicable to our contemporary world, which is following the first and second revolutions.

It is clear now that these revolutions represent a succession of technological innovations that show no sign of stopping and are speeding up to the point where some speculate that a limit may be reached, a singularity, an event horizon that we not only find difficult to grasp but impossible to predict (Kurzweil, 2005). One of the consequences of these waves of change which weave in and out of one another is that they are increasing the complexity of cities and society in ways that make it harder and harder to trace their impacts on these systems. Just as cities of the remote past were physically built up layer upon layer as 'tells', the cities of this century are being moulded by layer upon layer of new information technologies that quickly morph into one another, making it ever more difficult to dissect them into their component parts.

Into this maelstrom of technologies, the city stands out as a critical focus. Prior to the first Industrial Revolution, cities were the exception rather than the rule. Even in Britain at the beginning of the first revolution around 1750, the percentage of persons living in cities was only about 20%. 250 years later this is now over 90% while the 10% not living in cities are very highly urbanised, involved with city-like activities although living outside of cities in terms of their physical location. Worldwide those living in urban areas are now nearing 60% and by the end of this century, most will be living in cities of one size or another (Batty, 2011). From this perspective, the concept of the city appears somewhat outdated and problematic, largely because one of the most widely agreed definitions of a city is physical; cities in the past have tended to have a hard edge which is largely based on density but it is immediately obvious that such boundaries are highly blurred. Before the Industrial Revolution and the development of new transportation technologies such as rail and car, the hard edge was often deeply impressed as a fortress wall to keep the enemy out (and often to keep the population in) and there was much less ambiguity in defining this separation from the wider rural and agricultural hinterland in which a town or city was located.

From this discussion of a clean break with the pre-industrial past, we present three speculations about the urban future which we will discuss in turn.

Speculation I: The concept of the city in an entirely urban future will disappear

It is arguable that the concept of the city will be increasingly regarded as an artefact from an earlier age. As the digital revolution continues to gather pace, we will all be equipped with a means of contacting any and

everyone else using the digital skin that is fast being established around the planet. By the end of this century if not before, there will be complete globalisation in terms of contact patterns. There may be local reactions against such globalisation but the advance of digital technologies is unlikely to be slowed, and it is possible that we will adapt our global networks to deal with quite legitimate local and parochial concerns that continue to be necessary to urban life.

As we adapt to faster and closer technological revolutions, the sharp break or discontinuity occasioned by the first Industrial Revolution will come into clearer focus where the concept of the city is central to our understanding. So our first speculation is that the concept of the city will disappear, if not in name, for there will still be as much segregation and clustering in urban areas as before, but in terms of function. In fact the term ‘city’ is likely to be subsumed into the term ‘urban’ although urban is a concept that in turn is focused on the process of urbanisation, which in general is used to define the way populations cluster in dense areas which are towns and cities. In many definitions of what it means to be urban and what processes characterise those of urbanisation, there is an intrinsic circularity between the terms ‘city’ and ‘urban’. In this context, there are two countervailing forces which dominate the growth of cities, the first being total world population, which provides the envelope for different sizes of city, and the second being the distribution of these different sizes, which generally follow a rule that pertains to the fact that as cities grow, the number of larger cities decreases; that is, there are many more small cities than big cities determined by the fact that resources are limited, and not all cities can grow to be big.

In a world which is composed only of cities, the distribution of different sizes will sum to the world’s population, which we can define as P . The distribution of city populations can be ordered by first ranking them as $P_1 > P_2 > P_3 \dots > P_R$ and then noting that the widely accepted model of these frequencies called the rank-size rule or Zipf’s Law can be stated as $P_r \propto r^{-1}$ where P_r is the population of the city ranked as r . We speculate that this model will continue to apply when the entire system – the world – is composed of cities and we can dimension this so that the smallest city is fixed as one unit of population $P_R \propto R^{-1} \propto 1$. We can use this to scale the distribution as $P_r = Rr^{-1}$ where the population of the largest city is $P_1 = R$ and the total population can be calculated as $P = R \sum_{r=1}^R r^{-1}$. This relationship is independent of the actual total population in the system and it can be further scaled to ensure that it is appropriately dimensioned. It does, however, depend on the number of cities in the system, which is R , and thus any scaling must be a trade-off between the form of the power law relationship, its parameter, and the actual sizes of the largest and smallest cities (Batty, 2021).

The last element in our first speculation relates to the fact that as cities have grown bigger and as they have begun to fuse, they have been categorised using historical definitions that pertain to different kinds of ‘polis’. Doxiadis (1968) defined an elaborate sequence from the smallest settlements – hamlet, village, . . . , polis, . . . to the largest – metropolis, megalopolis . . . ecumenopolis (Batty, 2018). But the most widely used contemporary term is ‘megalopolis’, first used by Patrick Geddes but somewhat generically defined by Gottman (1961) as a ‘. . . very large city – an extended urban area with millions of people’ which he suggested could be the north-eastern seaboard of the United States. In fact, since Gottman’s work, the term ‘mega city’ has come to be used for really big agglomerations such as the Greater Bay Area connecting Hong Kong, Shenzhen and Guangzhou with Zuhai and Macao, the Greater Beijing and Shanghai regions, while elsewhere, much of north-western Europe falls into this category of polycentric development, and of course the north-eastern seaboard of the USA.

One last point to qualify our first speculation involves the fact that there are likely to be limits on the growth of future cities which we do not yet know. Although cities are getting bigger, relatively the biggest cities are showing a relative fall in their size compared to all cities. As cities fuse together, the resultant agglomerations function very differently from individual core cities and it is likely that the qualitative changes in size, scale and movement which occur when cities get bigger will be transomed in ways that we are not yet able to guess. This is due to the fact that there are still quite severe limits on how far we can move during the working day, but what is clear is that different kinds of movement are likely to conspire to change the basis for physical location in a globally connected world.

Speculation 2: As cities get bigger, they change qualitatively

The biggest cities throughout history are qualitatively different from smaller cities, but to be a big city you must be a small city first. This means that at any point there are subtle transformations taking place that enable their populations to engage and interact in different ways as they grow. The most obvious change is in the pattern of potential encounters that can take place. In a city with a population P , there are total of P^2 potential links between individuals and as the city grows this number explodes exponentially. This is sometimes referred to as Metcalfe's Law, which relates to the power of a network. In fact in human systems, most of these potential linkages cannot be exploited – there are too many – and it is much more likely the number of contacts grows to a power a lot less than 2. This still generates growth of contacts Y super-linearly, but it is much less than the total potential, which is much more like $Y = kP^{1.2}$, as has been demonstrated for various urban attributes such as income by Bettencourt et al. (2007) amongst others.

What this implies is that as cities grow they generate economies of scale, 'agglomeration economies' as Alfred Marshall coined them at the end of the nineteenth century. The formal relationships implied by this growth are part of the science of allometry, which is widely applied to the changes in the size and metabolism of biological systems. Applied to cities, then, attributes such as income scale super-linearly with population, implying that if one wishes to increase income per capita, one must grow the city. This argument runs deep in fact, for it suggests that we should all aim to live in bigger cities if we wish to generate a greater personal prosperity. Big cities are thus cool in the vernacular of the young. However, so far there is little sense of what happens if there were strict limits on city size, but in a world where total population is capacitated, then these relationships are likely to break down. However, the world that is emerging here is one where increasingly the notion of a city with a fixed physical size no longer holds but is being rapidly being replaced with the idea of a city being a function of its networks. In this context, economies of scale are more likely to hold with respect to the size of networks that individuals are able to activate. This suggests that such economies might only be detectable if cities are broken down into groups that are associated with different kinds of networks. The notion of relating population to its key attributes at a much more disaggregate level than in terms of whole cities is entirely consistent with the idea that the urban future will be based on cities that fuse into one another and no longer have the coherence of separate agglomerations with their own internal functioning.

There is another dimension to this kind of scaling. Physical infrastructures tend to scale sub-linearly with city size, but this does not imply that these are generating diseconomies of scale. The area given over to transport, for example, scales in this way, implying that this kind of space is used more efficiently as cities grow in size; in short, cities need less of it. Although our second speculation will need to be qualified as the urban future becomes one which is populated by networks of many kinds, and as technological innovations are made over much wider areas than the physical extent of any city itself, the notion of qualitative change due to size and scale will continue to be centre stage in this new future.

Speculation 3: Cities function in time as well as space

Many processes that determine how cities function and how we function within them are temporally dynamic. However, hitherto our classification of cities tends to have been mainly spatial with respect to their size, building on the 'polis' definitions that are associated with scholars such as Doxiadis who we noted above. Most objects that exist in space and time, however, are capacitated in some way and our first speculation pertaining to the growth of world population is a straightforward application of the way an activity grows to fill a capacitated system. This is usually modelled using the logistic function that is a capacitated exponential. This simulates the change in population dP at a time instant dt from $dP/dt = \beta P(1 - P/K)$ where P is the total population at that instant, β is the growth rate and K is the capacity limit, that is, the total population at the steady state which cannot be exceeded. It is easy to see how this is accomplished because the two terms on the RHS of the logistic equation moderate each other; the first term is the exponential and the

second the damping term that relates to how close the total population is to the limit. The total population $P(t)$ at time t can be derived formally as $P(t) = KP_0 \exp(\beta t) / [K + P_0 (\exp(\beta t) - 1)]$, where P_0 is the starting population.

In fact there are many processes like this in cities that determine how activities grow and change, but these are most relevant to low-frequency events such as the growth of populations rather than higher-frequency activities such as those that pertain to processes involving fine-scale mobility. The complexity of this picture can easily be visualised as many of these kinds of logistic processes operating in parallel and related to one another in both straightforward and complex ways involve networks which link elements of these processes to one another. If we expand the temporal spectrum from eras to centuries then there are many processes that need to be modelled in the way we illustrated in our first speculation. The clean break or great transition that is implied in our first speculation is highly relevant to the way the future city might emerge. So far we have assumed capacitated growth, but in an intrinsically unpredictable world it is not clear that world population will converge to a steady state this century. It might oscillate if capacity limits are breached and invoked in random fashion, generating scenarios for the future that are more like the dire outcomes that the Club of Rome forecast some 50 years ago (Meadows et al., 1972).

In terms of dynamics, change and movement at very fine time scales exist at the opposite end of the temporal spectrum. Smart city technologies are continually being embedded as smart sensors into the built environment, providing massive streams of data in real time with respect to second-by-second operation of the city across different domains. This data ultimately turns into long-term data which is consistent with lower frequencies of change, but so far we do not have a clear set of theories that direct us in how to use much of this data other than in the performance monitoring that such data is designed to reveal. Much of it remains as an information exhaust where its structure needs to be extracted using various machine learning methods.

Articulating these processes in terms of flow equations is, however, possible and there are many transportation models that are focused on this, but we do not yet have a comprehensive perspective on how all of this might fit together. In extending our speculations, then, it is likely that network processes will become ever more significant in our thinking about the shape of future cities. The great change from pre- to post-industrial is not only defined by a transition from city to urban or from few cities to many but also from a world where location is a dominant feature of cities to one where networks dominate – from actions to interactions. These networks that will define the future city in so far as they exist at all in physical terms lie in clusters, strings, lattices and other topological structures that are no longer adjacent to one another in terms of our traditional spatial dimensions. The challenge is thus to build an urban science that provides robust coherence to our understanding of the shape of future city.

Declaration of conflicting interests

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author received no financial support for the research, authorship, and/or publication of this article.

References

- Batty M (2011) Commentary: When all the world's a city. *Environment and Planning A* 43(4): 765–772.
- Batty M (2018) *Inventing Future Cities*. Cambridge, MA: The MIT Press.
- Batty M (2021) The size of cities. In Glaeser E, Kourtit K and Nijkamp P (eds) *Urban Empires: Cities as Global Rulers in the New Urban World*. London: Routledge, 210–228.
- Bettencourt LMA, Lobo J, Helbing D, Kuhnert C and West GB (2007) Growth, innovation, scaling, and the pace of life in cities. *Proceedings of the National Academy of Sciences* 104(17): 7301–7306.

- Bell D (1974) *The Coming of Post-Industrial Society*. New York: Harper Colophon Books.
- Doxiadis CA (1968) *Ekistics: An Introduction to the Science of Human Settlements*. London: Oxford University Press.
- Gottman J (1961) *Megalopolis: The Urbanized Northeastern Seaboard of the United States*. New York: The Twentieth Century Fund.
- Kurzweil R (2005) *The Singularity Is Near: When Humans Transcend Biology*. New York: Viking.
- Meadows DH, Meadows DL, Randers J and Behrens WW (1972) *The Limits to Growth: A Report for the Club of Rome's Project on the Predicament of Mankind*. New York: Potomac Associates, Universe Books.
- Schwab K (2016) *The Fourth Industrial Revolution*. New York: Portfolio Penguin.
- Toffler A (1980) *The Third Wave*. New York: William Morrow.
- Wiki (2022) Estimates of historical world population. Available at: https://en.wikipedia.org/wiki/Estimates_of_historical_world_population (accessed 4 July 2022).

Michael Batty FRS FBA is Bartlett Professor of Planning at University College London where he is Chair of the Centre for Advanced Spatial Analysis (CASA) and a Turing Fellow in the Alan Turing Institute. With Wei Yang, in 2021, he led the Taskforce for Digital Planning, whose report focuses on the need for British planning to renew itself in adopting state-of-the-art digital technologies, see <http://www.digital4planning.com/>.