Modeling megacity futures

Dr. Paul M. Torrens, Geosimulation Laboratory, School of Geographical Sciences, Arizona State University, Tempe, AZ 85287-5303, USA. Email: torrens at geosimulation dot com.

Symposium on Megacities, Sonny Astani Department of Civil and Environmental Engineering, Viterbi School of Engineering, University of Southern California, Los Angeles, November 10-11, 2008.

The significance of megacities

The rationale for studying urban systems and phenomena is varied and compelling. Urban activity is among the most significant of the Earth's land-uses. Cities host vast amounts of the world's built and technical infrastructure, they have served among the most important engines of land cover change through history, and they are significant sources of anthropogenic contributions to the Earth's climactic systems. Cities also serve as hubs of human activity: they provide the ambient human infrastructure for much of the world's social, economic, and cultural systems, as well as providing the substrate that houses the majority of the world's population. Urban systems are still growing in extent and volume throughout the world. In many areas, the pace of urban expansion is actually accelerating, sometimes at unprecedented rates. This is particularly true in the world's megacities: unified urban agglomerations with populations of at least ten million inhabitants. It is here, in megacities, that the greatest engines of the world's urban activity—and all of its associated problems and promise—are to be found.

In the last thirty years, the number of megacities in the world has increased from three to twenty. The geography of this mega-urbanization is uneven. Most megacities in the developed world are projected to reach a level of stasis in their growth, growing at slower rates as their populations saturate their urban environment and the role that they play in their constituent national systems—and globally—cements, at least for the time being. Growth in the Los Angeles-Long Beach-Santa Ana megacity is forecast to expand by only 6.5% (+0.8 million, to 13.1 million total) between 2005 and 2015, while that of Tokyo is set to appreciate by <1% (+0.3 million, to 35.5 million total) over the same time-frame (Moore and Gardner 2007). No net growth is projected for the Osaka-Kobe megacity over that period (its population is set to remain steadfast at 11.3 million in total) (Moore and Gardner 2007). Meanwhile, megacities in the developing world are forecast to accelerate in their growth: Lagos megacity is projected to expand by 48%, adding 5.2 million people (to 16.1 million total) between 2005 and 2015, Dhaka is estimated to grow by 35% (+4.4 million, to 16.8 million total), Karachi by 31% (+3.6 million, to 15.2 million total), Jakarta by 27% (+3.6 million, to 16.8 million total), and Kolkata by 19% (+2.7 million, to 17 million total) over the same time period (Moore and Gardner 2007). As megacities grow and consolidate with massive tangible footprints and huge populations, so also will their influence on the world's physical, natural, social, and technical systems expand and intensify. The pace of their emergence, development, and growth has, to a certain extent, outpaced our ability—as scientists—to keep track of their driving mechanics. Appreciating and understanding the future evolution of megacities is critical in explaining the futures of the world's demography, economic markets, climate variability, innovation, and in postulating about many other factors.

Megacities are, however, a difficult phenomenon to study. Their enormous size veils many of their attributes to inquiry. In cases, their emergence is a relatively recent phenomenon. The developing Guangzhou megacity in China, for example, had a population of just 2.7 million thirty years ago; by 2015 it is expected to reach 10.4 million (Moore and Gardner 2007). Although it is now home to 11 million people, the Lagos megacity in Nigeria had a population of 1.9 million thirty years ago (Moore and Gardner 2007). Experimenting with such mammoth, unwieldy, and rapidly adapting systems in any sort of tangible fashion on the ground is understandably prohibitively difficult.

Urban simulation as a looking-glass for megacity futures

In circumnavigating the difficulties of studying megacities, we may turn to simulation as an alternative—artificial—laboratory for experimenting and theorizing about their present conditions, as well as their past and future trajectories.

A plethora of models exist for simulating urban *sub-systems* at macro-scales (e.g., inter-regional migration (McHugh and Gober 1992), scaling and allometry in global city-size distributions (Batty 2008), and the geography of national urban agglomeration economies (Fujita *et al.* 2001; Krugman 1996)), as well as characterizing sub-systems at meso-scales (e.g., intra-urban traffic flow (Nagel and Schreckenberg 1992), formation of urban heat islands (Brazel *et al.* 2000), and urban epidemic dynamics (Eubank *et al.* 2004)), and micro-scales (e.g., pedestrian activity along city streets (Haklay *et al.* 2001), vehicle parking behavior (Benenson *et al.* 2006), and emergency evacuation behavior (Nara and Torrens 2007)). Such sub-systems are found in megacities, but sub-system models do not address megacities as entities in their own right and seldomly consider the dynamics of these systems as special cases in megacity contexts. Global climate models, for example, often treat cities as a simple binary classification—urban or not urban—in accounting for land cover in the boundary-layer of the Earth's climate systems.

In an ideal situation, we could couple many sub-system models together to generate a *super-model* that explains the intricate inner-workings of megacities at the detail of its constituent components, but this is insufferably difficult to achieve in practice because individual models are often developed for independent purposes, with purpose-specific data models, methodological approaches, spatial resolutions, constraining assumptions, system closures, time-scales, and so on. Several meta-models have attempted to couple the information flow between diverse sub-system models—often designed as "stocks and flows" models that determine the elasticity in intra-system relationships (Forrester 1969)—but these are not strictly integrated models.

The closest analog to the ideal of a super-model is the large-scale urban model, most commonly developed for operational use by metropolitan planning organizations in estimating the ability of the city to provide future urban services. Previous generations of these models were developed as coupled land-use and transportation simulators, in which a land-use model generated trips which were then simulated over large-geography urban systems as traffic flows; the DRAM/EMPAL model was a widely-deployed example (Putman 1983). These models were often built on microsimulation (Clarke 1996) and regional science (Isard 1975) methodologies, which for the most part used relatively simple heuristics (often based on physics of spatial interactions (Fotheringham and O'Kelly 1989) and parametric statistics for estimating discrete choices for activity types and locations (Louviere *et al.* 2000)) to extrapolate future values from coarse-resolution socioeconomic data-sets. Large-scale urban models have been applied in evaluating planning and policy alternatives in many megacities, with mixed usefulness (Batty 1994). In traditional form, large-scale models tend to treat land-use and transport dynamics in a crude, abstract fashion, and although they may be loose-coupled to environmental models in some cases (tying traffic flows to aggregate emissions estimates, for example (Wegener 1994)), they largely fail to treat the full range of sub-systems that account for megacity formation and adaptation in sufficient detail.

More recently, a next generation of large-scale urban models has been developed as planning support systems, which more closely approach the ideal of an integrated, detailed model of an entire city-system. Detail has been added to these systems, in large part, by developing a slew of sub-models that handle demographics, lifecycle transition, migration patterns, diverse modes of transportation, land-use change, and property markets. Two planning support systems stand out in particular—the California Urban Futures models (Landis 2001), and UrbanSim (Waddell 2002). Urbansim, in particular, is relatively widely used in operational city planning. It was initially developed as an urban economic model, designed to estimate the future trajectories of urban land markets, but several efforts are underway to extend the model by integrating it with detailed activity, travel, and traffic models, as well models of urban natural environments, and the lifecycle of resource use in constructing large city-systems (Li *et al.* 2007).

A parallel thread of model development has been carried out in the sustainability sciences, focused predominantly on modeling the role of human-environment interactions upon land-use and land cover change. The sustainability science community has benefitted greatly from increasing availability of remotely-sensed data at finer-grained resolutions and covering longitudinal periods of time. Many of these models are focused on large city-systems, with an emphasis on the extension of expanding cities into the urban-rural interface through suburban sprawl, edge city formation, and exurban development (Parker *et al.* 2003).

The challenge of modeling complex adaptive systems

There are fundamental challenges in representing megacity systems in simulation, however. These massive urban behemoths are complex systems composed of many interacting sub-systems, each intertwined through a bewildering array of non-linear and dynamic phenomena that scale up and down and weave throughout the fabric of the past, present, and future. In the case of these huge urban agglomerations, the orders of magnitude in scaling from the individual to the system are many times greater than one would usually encounter. The number of state descriptors and linkages required to explain the functioning of megacities are also substantially greater, as are the potential trajectories for the system's state-space over time. Pinpointing the emergence of novel patterns and phenomena in megacity evolution is an arduous task considering the cacophony of actions and interactions within the megacity that must be scrutinized in order to identify such innovation. Traditional toolkits and methods for modeling cities and city-systems are limited in their ability to treat the complexities that drive urbanization on a mega-level. Those toolkits largely dictate the sets of questions that can be posed in simulation, constrained within the limitations of the specific assumptions that they make, rather than being flexible in handling multiply-interacting city systems across a variety of scales (Batty and Torrens 2005).

In response to these challenges, many model-developers have turned to complexity studies in search of methodology for treating the complexities inherent in large city-systems. Most have been built around automata (cellular automata, agent-based models, individual-based models, multi-agent systems, geographic automata) (Benenson and Torrens 2004), following the success of such tools in generating signature complexities in the Artificial Life community, economics, mathematics sciences, ecology, computer science, and physics (Wolfram 2002).

Much of the development work and application of complexity models to urban systems has been carried out in the geographical sciences. Many automata-based models have focused on some of the spatial drivers of megacity formation. In some cases, treatment of megacities in such models is explicit, for example, in modeling the formation of megalopolitan structures—massively cohesive city-systems (Gottmann 1967)—from the bottom-up using synthetic developer-settler agents (Torrens 2006). In other instances, drivers of megacity development are represented less explicitly, as agents of fringe urbanization and forces of urban polycentricity, or the factors that determine the allometry of urban central place hierarchies, for example (Batty 2005).

For the most part, however, development of these approaches is in very early stages and few models make it out of sheltered laboratory settings and into use on the ground or traverse the journey necessary to span disciplinary gaps. Algorithmically, the models generally retain an overarching focus on simple rule-of-thumb heuristics from urban studies (grow on the edge of the urban mass, don't build on steep slopes, fill-in interstitial urban sites if they are surrounded by sufficient development, and so on (Clarke *et al.* 1997)). In other cases the models are largely data-driven: their algorithms focusing on spatially distributing the data that is fed to them; these models are generally only as good as the data that they are fed and little reliance can be placed upon their future extrapolations. They are, as the cliché goes, "tools to think with" (Negroponte 1995) rather than serving as decision support systems.

Nonetheless, significant developments have been made in advancing the state-of-the-art in modeling urban sub-systems and this points to the potential of complexity-oriented approaches as a candidate for ideal super-models of megacities. Much of the innovation has been achieved in building models of urban-scale traffic systems, at the level of individual drivers, their decisions, actions, and interactions, propagated up-scale and down-scale between the city and the road (Barrett *et al.* 1999; Torrens 2005). Sophisticated models of property markets and residential formation at the scale of interacting households and communities have also been built (Benenson *et al.* 2002; O'Sullivan 2002; Torrens and Nara 2007). In these instances, the path from individual agent to larger-scale system (and back again), and all of the complex interactions that take place in between can be relatively easily identified and expressed algorithmically, within spatial, temporal, and system confines, largely because there are long traditions of social science, behavioral, and economic research into these sub-systems, and data for calibrating and verifying models against ground-truth are often available. These remain, however, limited cases and they explain only one of many (isolated) components that drive the development and day-to-day functioning of megacities.

Challenges in modeling megacities

Perhaps the greatest challenge in building more useful models of megacities as artificial laboratories is the sheer size and complexity of the urban systems that they encapsulate. To a certain extent, abstract models that at least represent megacities in a proxy fashion should be useful as experimental toolkits, but other challenges remain.

In other fields (climatology, cosmology, macroeconomics, for example), standard models have been in place for many decades and these serve as a foundation for innovation in their respective scientific communities. There is no (robust) standard model for cities or megacities, largely because each city is rather unique in its composite patterns and processes in a much more variable way than structures in climatology, cosmology, and macroeconomics might be considered. Invariably, then, model-builders must start anew in constructing new tools and this slows the pace of innovation. Building a common platform for urban simulation, one that treats some of the more generic components of city-systems, may help to ease this constraint.

Sophistication in urban simulation is almost always closely allied to the availability of data, and the plentitude of data at high spatial and temporal resolutions, covering a multitude of urban sub-systems. With the exception of remotely-sensed imagery relating to land-cover, such data are often in short supply, particularly at the micro-scale. Recent developments in cyberinfrastructure for automated sensing and data collection over distributed sensor webs suggest that issues of data availability may be resolved in the near future, but data will most likely be in short supply for many urban sub-systems, particularly those relating to human decision-making. These data are simply too difficult to collect over megacities or to infer, using cell-phone records or patterns of vehicle or currency mobility, for example (although attempts to do so have been made (González *et al.* 2008)).

For systems in which data may be available, they are often required in massive volumes to feed very data-hungry urban simulations. Similarly, the data that complex simulations output are often in volumes that are many times greater in size than the resources that are initially input. Sophisticated dataware are therefore needed to visualize inputs and outputs in a scientific fashion and to mine data for knowledge discovery and generation. There has been a fantastic amount of innovation in visualizating complex information, in the development of information systems for handling massive data-sets, and in crafting intelligent routines for knowledge discovery, data-mining, and reality-mining of large data reservoirs. With few examples (Batty *et al.* 2001), much of this innovation has not yet been introduced to urban simulation, particularly as a decision support system.

Issues of calibrating, validating, and verifying complex urban simulations often compound these problems. Because megacities are such large and unwieldy phenomena, garnering ground truth for the purposes of model-fitting is a very difficult task. Models are therefore often built blindly, as proofs-of-concept, or are built from theory, which is almost always anecdotal, qualitative, and even normative in nature. Building robust models on such a permeable foundation is quite a difficult undertaking. Improving data resources and related dataware may help to resolve such issues, but complicating factors remain as grand challenges, particularly in treating uncertainity and stochasticity in the interface between models, data, and “truth”.

Large-scale urban models, if simulated with any serious degree of fidelity to the mega-systems that they are tasked in representing, are usually massive software engineering projects that require considerable computing resources. To some degree, principles of encapsulation, abstraction, clustering, scheduling, and distributed computing from high-performance computing may be used to great advantage in urban simulation and already are, for example, in traffic modeling (Nagel and Rickert 2001), where road segments may be neatly parsed and passed between processing units on parallel systems. Considering megacities more comprehensively, however, involves treating massively dynamic and interacting agents and agencies with many-to-many relationships that scale-up, scale-down, and act and interact with complex and fluid feedback contingencies. Such processes and phenomena are not as easily and discretely packaged as computable packets.

Finally, and perhaps most importantly, megacities and their constituent drivers remain largely as unknowns in social science, behavioral science, and even in the design and engineering sciences. They are organic, bottom-up, dynamic, and adaptive systems that do not readily make sense microscopically, synoptically, or from vistas in between. The science of modeling megacities is being developed at the same time that theories are being forwarded, evaluated, accepted, rejected, and modified. Concurrently, megacities are growing, adapting, accelerating, and reaching equilibrium in different places and contexts. There is a need for a simulation methodology that can flexibly keep step with these developments on the ground and in our theoretical discussions. Automata-based approaches and related artificial intelligence methods are the most likely candidate to satisfy those requirements, but much work remains in developing that science to a condition of maturity before their promise can be more fully realized.

References

Barrett, C. L., R. J. Beckman, K. P. Berkbigler, K. R. Bisset, B. W. Bush, S. Eubank, J. M. Hurford, G. Konjevod, D. A. Kubicek, M. V. Marathe, J. D. Morgeson, M. Rickert, P. R. Romero, L. L. Smith, M. P. Speckman, P. L. Speckman, P. E. Stretz, G. L. Thayer, and M. D. Williams. 1999. TRANSIMS (TRansportation ANalysis SIMulation System). Volume 0: Overview. Los Alamos: Los Alamos National Laboratory.

Batty, M. 1994. A chronicle of scientific planning: the Anglo-American modeling experience. *Journal of the American Planning Association* 60 (1):7-16.

———. 2005. *Cities and Complexity: Understanding Cities with Cellular Automata, Agent-Based Models, and Fractals*. Cambridge, MA: The MIT Press.

———. 2008. The size, scale, and shape of cities. *Science* 319 (5864):769-771.

Batty, M., D. Chapman, S. Evans, M. Haklay, S. Kueppers, N. Shiode, A. Smith, and P. M. Torrens. 2001. Visualizing the city: communicating urban design to planners and decision-makers. In *Planning Support Systems in Practice: Integrating Geographic Information Systems, Models, and Visualization Tools*, eds. R. K. Brail and R. E. Klosterman, 405-443. Redlands, CA and New Brunswick, NJ: ESRI Press and Center for Urban Policy Research Press.

Batty, M., and P. M. Torrens. 2005. Modeling and prediction in a complex world. *Futures* 37 (7):745-766.

Benenson, I., S. Birfur, and V. Kharbash. 2006. Geographic Automata Systems and the OBEUS software for their implementation. In *Complex Artificial Environments*, ed. J. Portugali, 137-153. Berlin: Springer.

Benenson, I., I. Omer, and E. Hatna. 2002. Entity-based modeling of urban residential dynamics: the case of Yaffo, Tel Aviv. *Environment and Planning B: Planning and Design* 29:491- 512.

Benenson, I., and P. M. Torrens. 2004. *Geosimulation: Automata-Based Modeling of Urban Phenomena*. London: John Wiley & Sons.

Brazel, A., N. Selover, R. Vose, and G. Heisler. 2000. The tale of two climates—Baltimore and Phoenix urban LTER sites. *Climate Research* 15:123-135.

Clarke, G. P., ed. 1996. *Microsimulation for Urban and Regional Policy Analysis, European Research in Regional Science 6*. London: Pion.

Clarke, K. C., S. Hoppen, and L. Gaydos. 1997. A self-modifying cellular automaton model of historical urbanization in the San Francisco Bay area. *Environment and Planning B* 24:247-261.

Eubank, S., H. Guclu, A. Kumar, M. V. Marathe, A. Srinivasan, Z. Toroczkai, and N. Wang. 2004. Modelling disease outbreaks in realistic urban social networks. *Nature* 429 (6988):180-184.

Forrester, J. 1969. *Urban Dynamics*. Cambridge, MA: The MIT Press.

Fotheringham, A. S., and M. E. O'Kelly. 1989. *Spatial Interaction Models: Formulations and Applications, Studies in Operational Regional Science*. Dordrecht: Kluwer Academic Publishers.

Fujita, M., P. Krugman, and A. J. Venables. 2001. *The Spatial Economy: Cities, Regions, and International Trade*. Cambridge, MA: The MIT Press.

González, M. C., C. A. Hidalgo, and A.-L. Barabási. 2008. Understanding individual human mobility patterns. *Nature* 453 (7196):779-782.

Gottmann, J. 1967. *Megalopolis: The Urbanized Northeastern Seaboard of the United States*. Cambridge, MA: MIT Press.

Haklay, M., D. O'Sullivan, M. Thurstain-Goodwin, and T. Schelhorn. 2001. "So go downtown": simulating pedestrian movement in town centres. *Environment and Planning B* 28 (3):343-359.

Isard, W. 1975. *Introduction to Regional Science*. Englewood Cliffs, New Jersey: Prentice-Hall.

Krugman, P. 1996. *The Self-Organizing Economy*. Malden, MA: Blackwell.

Landis, J. 2001. CUF, CUF II, and CURBA: a family of spatially explicit urban growth and land-use policy simulation models. In *Planning Support Systems: Integrating Geogrphic Information Systems, Models, and Visualization Tools*, eds. R. K. Brail and R. E. Klosterman, 157-200. Redlands, CA: ESRI Press.

Li, K., Z. Peng, J. Crittenden, S. Guhathakurta, A. Sawhney, H. Fernando, P. McCartney, N. Grimm, H. Joshi, G. Konjevod, Y. Choi, S. Winter, D. Gerrity, R. Kahhat, Y. Chen, B. Allenby, and P. M. Torrens. 2007. Development of a framework for quantifying the environmental impacts of urban development and construction practices. *Environmental Science and Technology* 41 (14):5130-5136.

Louviere, J. J., D. A. Hensher, and J. D. Swatt. 2000. *Stated Choice Methods: Analysis and Application*. Cambridge: Cambridge University Press.

McHugh, K., and P. Gober. 1992. Short-term dynamics of the U.S. interstate migration system, 1980–1988. *Growth and Change* 23 (4):428-445.

Moore, S. K., and A. Gardner. 2007. Megacities by the numbers. *IEEE Spectrum* June:24-25.

Nagel, K., and M. Rickert. 2001. Parallel implementation of the TRANSIMS micro-simulation. *Parallel Computing* 27 (12):1611-1639.

Nagel, K., and M. Schreckenberg. 1992. A cellular automaton model for freeway traffic. *Journal de Physique I* 2 (12):2221-2229.

Nara, A., and P. M. Torrens. 2007. Spatial and temporal analysis of pedestrian egress behavior and efficiency. In *Association of Computing Machinery (ACM) Advances in Geographic Information Systems*, eds. H. Samet, C. Shahabi and M. Schneider, 284-287. New York: Association of Computing Machinery.

Negroponte, N. 1995. *Being Digital*. London: Coronet.

O'Sullivan, D. 2002. Toward micro-scale spatial modeling of gentrification. *Journal of Geographical Systems* 4 (3):251-274.

Parker, D. C., S. M. Manson, M. A. Janssen, M. J. Hoffmann, and P. Deadman. 2003. Multi-Agent System models for the simulation of land-use and land-cover change: a review. *Annals of the Association of American Geographers* 93 (2):314-337.

Putman, S. H. 1983. *Integrated Urban Models*: Pion.

Torrens, P. M. 2005. Geosimulation approaches to traffic modeling. In *Transport Geography and Spatial Systems*, eds. P. Stopher, K. Button, K. Haynes and D. Hensher, 549-565. London: Pergamon.

———. 2006. Simulating sprawl. *Annals of the Association of American Geographers* 96 (2):248-275.

Torrens, P. M., and A. Nara. 2007. Modeling gentrification dynamics: A hybrid appraoch. *Computers, Environment and Urban Systems* 31 (3):337-361.

Waddell, P. A. 2002. UrbanSim: modeling urban development for land use, transportation and environmental planning. *Journal of the American Planning Association* 68 (3):297-314.

Wegener, M. 1994. Operational urban models: state of the art. *Journal of the American Planning Association* 60:17-29.

Wolfram, S. 2002. *A New Kind of Science*. Champaign, IL: Wolfram Media, Inc.