Computational Urban Modeling: From Mainframes to Data Streams

Vahid Moosavi
Chair for Computer Aided Architectural Design (CAAD) and Future Cities Laboratory, ETH Zurich, 8092 Zurich, Switzerland
svm@arch.ethz.ch

Abstract

Assuming computational technologies as a dominant factor in forming new scientific methods during the last century, we review the field of computational urban modeling based on the ways different approaches deal with evolving computational and informational capacities. We claim that during the last few years, due to advancements in ubiquitous computing the flow of unstructured data streams have changed the landscape of empirical modeling and simulation. However, there is a conceptual mismatch between the state of the art in urban modeling paradigms and the capacities offered by these urban data streams. We discuss some alternative mathematical methodologies that introduce an abstraction from the traditional urban modeling methodologies.

Introduction

Historically, the way we communicate and conceive of our environments is dictated through the way we encode the target phenomena. In a scientific language, these encoded views to reality are the models we use to represent different phenomena. The first underlying idea in this work is that in each period of human history models and modeling approaches have been influenced by more abstract bodies of thinking or technologies of thought (Foucault 2002). In the realm of cities and city modeling, which is the focus of this paper, one can refer to city models such as the City of Faith, the City as a Machine and the Organic City (Lynch 1984) or to refer to models based on the underlying urban elements such as Enclave, Armature and Heterotopia (Shane 2005) and specially, since the advent of computers from the second half of the last century, as the age of informational cities (Castells 1989). Assuming the interrelations between technologies and the way we are looking at our cities, in this work we investigate the historical developments in computational technologies and urban modeling approaches in parallel. First we draw a historical map of computational capacities starting from mid 1950s. Next, we discuss the historical developments of different urban modeling approaches alongside the evolution of computational capacities. Finally, we claim that from the last few years a fundamental shift in the role of computation in our cities has happened, which has opened up a new way of looking at the city related phenomena. This shift demands an inversion in the concept of city modeling from a theoretical point of view. As a result, alternative classes of mathematical concepts are required as the backbone of new modeling and simulation methodologies.

In each section we quickly refer to main mathematical theories behind different well-known urban modeling approaches and in fact, this text aims to find an abstract understanding of computational modeling efforts and does not aim to be a comprehensive and detailed literature survey such as (Batty 2009 and Wilson 2012).

Generations in Modeling Capacities

In this section we present the main milestones in the history of computation by highlighting their offered capacities and functions. Each new capacity demands for a new way of looking at the concept of modeling.

Analytical Power

Descriptive Theories as the ground, Analysis, Static Models, Natural Models and Cultural Cosmos

Classically there are two general approaches for solving a mathematical model, known as *analytical* and *numerical* approaches. Although for lots of mathematical problems, there is no analytical solution, for a long time the manual calculation of numerical algorithms for large-scale problems was cumbersome and expensive. Therefore, historically the majority of modeling approaches had more tendencies toward finding general laws of nature in forms of descriptive theories, which could be generalized and applied analytically to other cases. We call the category of the models that are based on this analytical capacity as *natural models*.

Computing Power

Algorithms as the Ground, Numerical Methods, Simulation, Dynamic Models, Rational models

Looking at computers as *abstract machines (Hovestadt and Buhlmann 2013)*, given a logical algorithm (a code) to deal with a problem, a computer is able to execute this code in a way much faster than human can do that. The invention of the computer in mid 20th century rapidly increased another way of encoding of the real world phenomena through systematic numerical procedures. We call, this category of models as *rational models*. In a rational model the role of the modeler is to build up a consistent system of logical algorithms that can be executed by a numerical computer simulation to imitate a certain real world phenomena such as urban traffic, land use dynamics, economic activities and so on.

Historically, there have been different technologies of computation starting from main frames, and then to democratization of computing through personal computers and to microcomputers, which are still getting faster and more powerful in an exponential rate. However, regardless of the speed and power of computing machines, what we want to highlight here is the function of numerical computing compared to analytical power as two different capacities for scientific modeling.

Alongside the hardware advancements, computational technologies have been developing in the field of algorithm design or the so-called computer science. Depending on the encoding approach of the real world phenomena, there are different categories of computational modeling methods, which are more or less similar among different disciplines. We will discuss this issue in the next sections on analyzing the state of the art in urban modeling.

Computational Networks

Computation as the Ground, Microprocessors, Sensors, Mobile Phones, Internet, Emergence, Semantic Web and Rise of Structured Data

Alongside the developments of computing technologies, advancements in communication technologies gradually opened up another level, in which computing powers is given as the ground, while what is important is the communication between computing systems. Therefore, new phenomena such as network of sensors, mobile phones or computers and the Internet started to emerge in human societies. Gradually from this time, considering the amount of embedded systems in many real world applications, computers as "computing machines" became the ground to introduce a new function, emerging on top of computational networks. As a by-product of these network of computing and communicating machines, gradually the amount of digital data started to increase as well. Almost around the same time starting from mid 1990s, technical terms such *data mining* as the methodology to explore digital data (mainly structured data) started to be a hot topic among the modelers.

Data Streams

Data as the Ground, Ubiquitous and Pervasive Computing, Social Media, Smart Phones, Web 2.0, Mobile Apps, Crowd Sensing, Data Deluge, Unstructured Data, Complex Models

Gradually, with rapid advancements both in the level of computing power and the networks of computing systems and the rapid growth in social media, during the last few years we have encountered to a new stage, in which on top of ubiquitous computing and communicating systems as the ground, a new level of abstract phenomenon started to emerge, just like a picture of a smiling face on top of thousands of RGB pixels in an image, where each pixel is metaphorically a computing and communicating system.

During the last 10 years, we have begun to experience an exponential growth in the amount of information available, together with the mobile computing devices most people use on a daily basis. This is often called a data deluge. Next to the challenges these changes bring, we can also see how new areas for research and practice are emerging. To just mention a few, one can refer to Big Data, Data Science, Ubiquitous Computing (Greenfield 2006), Pervasive Sensing (Hansmann 2003), Reality Mining (Eagle and Pentland 2006), Citizen Science (Paulos et al. 2008), Social Network Analysis and Location Based Social Network Analysis (Jing et al. 2012).

It seems clear today that the classic paradigm of observation and data gathering has changed radically. Data is produced on an everyday basis, from nearly any activity we engage in, and accumulates from innumerous sources and formats such as text, image, GPS tracks, mobile phone traces, and many other social activities, into huge streams of information in digital code. These unstructured and continuous flows, which can be called *Urban Data Stream*, can be considered as a new urban 'infrastructure'.

This notion of data is opposed to its classical notion, which data is produced mainly as the result of designed experiments to support specific hypothetical models. These new data streams are the raw materials for further investigations and similar to computing power they are new capacities for modeling. As a result of this new plateau, we are challenged to learn the new ways to grasp this new richness.

As we discuss in the next sections, urban data streams induce an inversion in the paradigm of computational modeling from an algorithmic and rational models to a new level that we call *complex models*.

In order to sum up, the following diagram shows the main waves of modeling capacities as we described in this section.

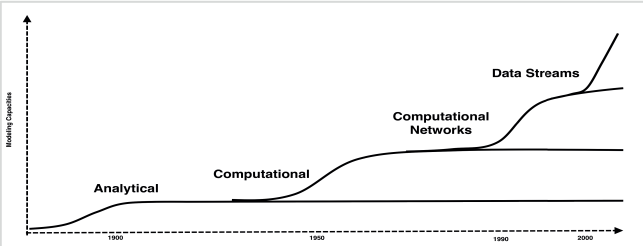

Figure 1- Historical trends of different modeling capacities

State of the Art in Computational Urban Modeling

Understanding the complexity, Theory Driven Models, Curse of Dimensionality, Complicatedness, Rationality, Idealization and Explicit Representation

There are several comprehensive literature surveys that review the field of urban modeling in a chronological manner such as (Batty 2009) or divided to spatial, temporal and functional scales (Wilson 2012). However, as we mentioned before in this work we are not going toward this direction, but we are looking for a way to compare the underlying concepts of the state of the art in urban modeling to the main functional capacities we derived in the last section. As figure 1 shows different capacities of modeling and data provision are not contradictory and in a way they are all coexisting. This means that at the current time some models are based on the analytical concepts of natural models or there are modeling efforts that are celebrating the concept of distributed computing for example.

The modeling approaches to be distinguished in a first class are based on analytical arguments. These approaches tend to regard the cities as a kind of “cultural cosmos” where they seek to identify statistical proportions, and the principle rules, which govern those proportions. May be the first model of this category is the agricultural land use model of Von Thunen in 1826 (Von Thunen 1966), which is based on an analytical model of the relationships between markets, production, and distance. Further we can refer to Christaller's Central Place Theory (1933), modeling the economic relations between cities and their hinterlands using geometric concepts. As one of the most famous approaches in this category one can refer to *Urban Scaling* (Batty 2005, 2008), in which the idea is to find proportions like the relation between city size and its energy consumption (Bettencourt, et al. 2007), or to finding universal laws of urban mobility patterns (Noulas, *et. al* 2012 and Simini, *et. al* 2012). Also, one can refer to *Space Syntax* method (Hillier 1984), which provides a set of network analytics based on the structure of the given street network and some universal assumptions about how people move in the streets. Due to availability of urban data in recent years, this category of urban modeling that sometimes called, *City Science*, is getting more attentions, but as they are based on analytical capacities (even though using computers and data), they result into *static* and *aggregate models*.

The modeling approaches that can be distinguished in a second class started to boom after the advent of computing power in the mid-1950s, and especially also since the introduction of so-called main frame simulations. Approaches, which mimic the behavior of another system, and try to optimize within the analogy they take as a preset, is for example *Urban Metabolism* (Wolman1965), which was based on an analogy of a city to a biological system. This analogy of a city as a biological organism became very popular with the advents of Cybernetics. But there are also examples, which import their analogies from other fields, like the approach of *Urban Dynamics* (Forester 1969), which established an analogy to models used in socio-economic systems, and from hydraulic systems in economy. It was introduced in the late 60ies and is known today as *System Dynamics* modeling approach. As another example, which became popular in the 60ies we can mention the idea of *Social Physics*, based on generalizations transferred from the realm of classical physics to social systems (Harris1964). For example, Rand Corporation was developing models to map the land-use dynamics, which were based on Newton’s theory of gravity; it proposed a system of equations to describe the different forces among urban actors (Lowry 1964). As another approach, one could refer to fractal based simulation models (Batty 2005), which apply recursive principles from *fractal geometry* and conceptions of self-similarity. As other similar theories, one can refer to applications of *Chaos theory*, *Catastrophe theory* and the *Bifurcation* under the umbrella of complexity theory (Batty 2005).

So far, these approaches mainly transfer overall generalizations from mathematical models and equation systems that are proved or well established in other fields like physics or economy.

From 1980s, at the same time that the concept of computational networks was booming and the relatively easier access to digital data in urban environments, there was a shift in the paradigms of urban modeling from *centralized models*, to *distributed models*.

This category of distributed urban modeling has caused a partially significant paradigm change in terms of methodology, from equation based models that are to represent the logics of overall system to distributed frameworks. Here, what is being modeled is the behavior of the components of such systems as individuals behaving differently over time (like humans as agents or cells in the land use models). Two main branches of micro-simulation approaches are those of *cellular automata* (Tobler, 1979), and *multi-agent systems* (Waddell and Ulfarsson 2004). Note that these agent-based approaches are mainly less developed in terms of *agent based learning*, which is a topic in the field of machine learning.

However, the modeling of the agents also requires the framework of a specific model or rule set, and this limits the capacities of these approaches – even if they are very sophisticated in many regards – in the same way as large-scale models are limited. Further, there are fundamental limits to these approaches in general. To just mention a few, one can refer to the *curse of dimensionality* problem (Bellman 1961), which means in principle these approaches of algorithmic modeling reaches to a limit in dealing with the complex phenomena such as cities, in which by adding one more realistic aspect of the target phenomena into the model, the demand for the computation and the data to tune these models increase exponentially. Hence they import simplifying assumptions in their presets, without being able to consider them comparatively and critically. As a result, the efforts of such modeling always became either very complicated or very simple and not sufficiently distinct for describing the complexity of urban environments with adequate benefit. For example, in one of the famous critiques to this class of urban models, one can refer to the paper Requiem for Large-scale Models (Lee 1973) which enumerates the particular limits of large scale urban models as the "Seven Sins of Large-Scale Models": 1) Hyper Comprehensiveness (attempt to explain too much with too many constraints and relationships), 2) Grossness (reliance on aggregate input), 3) Mechanicalness, 4) Expensiveness (high price of data and parameter estimates), 5) Hungriness (tremendous data requirements), 6) Tuningness (Tuning of the model until outputs conform to reasonable expectations), 7) Complicatedness (inability of the modelers to adequately understand their own creatures).

Transition: An Inversion in the Concept of Urban Modeling

Encapsulation of Complexity, Models in Coexistence with Data Streams, Data Literacy, Probabilistic Programing

The main hypothesis of this paper is that from the second half of the last century computational technologies are the dominant factors of the scientific modeling. However, as we discussed in the last section, the majority of traditional urban modeling approaches only utilize the numerical computing and simulation capacities (either centralized or decentralized) that started from mid 1950s. In these traditional modeling and simulation approaches prior knowledge in the form of idealized theories of the target phenomena is the primary element of the models, while data is to either validate the output of these frameworks or to tune the parameters of a given structural model. As a result, their use of new emergent capacities such as ever growing data streams is fundamentally marginal. The hypothesis is that in order to grasp the richness of the ever-growing urban data streams, an inversion in the concept of modeling is required. To better explain this conceptual inversion, the following example might be helpful. The left picture in figure 2 is the result achieved by Space Syntax method (Hillier 2009) applied into the street network of London. It is based on the rational assumption that the attractiveness, or the importance of certain street segments or city segments, can be regarded as a function of the physical connectivity in the urban networks. Therefore, computer power is needed to just compute these structural indicators of a given network. Further, the empirical data can be used to validate the results of these structural assumptions about patterns of movement. On the other hand, in figure 2 (right), only by visualization of available GPS tracks of taxicabs in Beijing (Jing et al. 2010), we already know the important segments of an urban network. This new capability of harnessing data streams is relatively new, and we think as a new emergent capacity it can be used as a starting point for development of a new modeling approach, beyond the current rational modeling approaches.

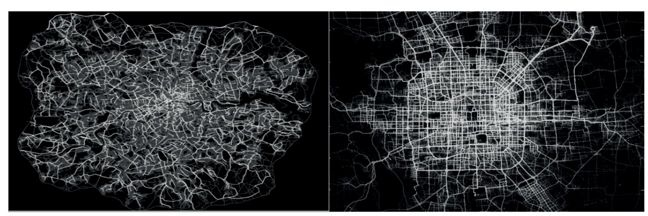

Figure 2- The Inversion in the concept of modeling: Visualizing the important urban segments based on logical assumptions (Hillier 1984) in left and in right, using GPS trajectory of taxicabs as an available urban data stream (Jing et al. 2010)

The underlying concept behind the traditional simulation models is to construct a logical set up to *understand the complexity* through its underlying mechanism. Then, by simulation one can imitate the behavior of the target phenomena. However, as we mentioned, this approach has fundamental limits. On the other hand, where we start by data streams is in an inverted situation. In fact, the data-driven illustration of GPS tracks of taxies in Beijing (assuming data is enough) *encapsulates the complexity* of traffic networks and is able to answer the questions of type "What" about this traffic network with a high degree of accuracy, without knowing "Why" these patterns are happening. Availability of data streams has promoted these generic emerging applications in many other similar cases.

For example, Google traffic live service is an aggregation of GPS data being emitted by people who are commuting along the streets and other sensory data. In the same manner, Google traffic live does not start with the assumptions whether people who are commuting take the shortest path or not, but it is simply a good render of what happens in the reality. As another interesting case, one can refer to Livehood project (Cranshaw, et al. 2012) that presents a methodology for identification and studying dynamics of land-use patterns in cities using data from social media and clustering methods from machine learning. Regardless of the amount of available public data, the methodology behind this project is in an opposite direction to classical land use models, where sets of rules of mobility and urban forms are presumed as the underlying mechanisms of defining a neighborhood. In a similar direction, one can refer to similar projects such as the applications of Endomondo mobile App data, in which collections of frequent running and walking patterns of people collected via this mobile app, can be used in urban design projects such as identifying emergent walk ways and to design better infrastructures for pedestrian movements.

However, one of the true concerns regarding these kinds of applications is that they might just end up to info-graphics or simple data analytics with no further functionalities. However, similar to the turn from analytical models to computational models, and the way we mastered computational algorithms we believe new paradigms of modeling is needed to be able to grasp the richness of these urban data streams. We call this set of future skills as *data literacy*.

Alternative Mathematical Modeling Concepts

Relationality, Observation Dependent Representation, Markov, Bayes, Self Organizing Map, Structural Learning

There are detailed discussions around the ideas of *representation* and *idealization* in the philosophy of scientific modeling (Weisberg 2007) as well as on the limits of set theoretical representation of models in terms of *abstract universals* and opportunities of using *concrete universals* from *category theory* to abstract from the current state of the art in scientific modeling (Ellerman 1988, Buhlmann 2013 and Moosavi 2014), which goes out of the scope of this work.

In a technical level, figure 3 shows two general approaches of representation, which lie at the hart of modeling approaches. Each circle in these diagrams stands for an object. These objects are symbolical, which means that they can stand for anything – be it people, cars, companies, buildings, streets, neighborhoods, cities, webpages in the internet, protein networks, networks of words in corpus of texts, or people and their activities in a social network. In the traditional (rational) modeling the first step is to define an ideal representation of the target phenomena, which as it is shown in figure 3 (left), this leads to a set of selected features of the real objects (for example, structure of the street network in Space Syntax method). Therefore, concrete instances of the object are assumed to be independent to each other and they are all compared indirectly by an abstract class definition, which acts as an external reference. On the other hand, in the right diagram there is no explicit and external reference system and as it is shown the identity of a single object is fully dependent on its relations to the other objects.

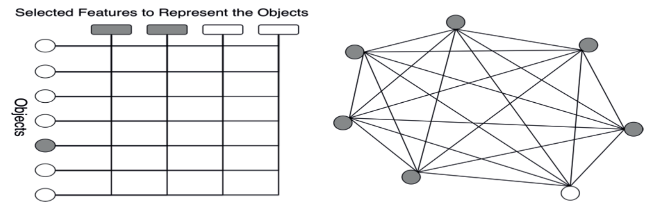

Figure 3 – (Left): Feature based representation of objects (rational model) and (right): Observation dependent representation of objects (Relational model)

In other words, if we take each object as a dimension (A feature), we can represent each object by the other objects, which are all in the same concrete level with no explicit need for ideal representation of the objects in an abstract level. In this way of representation and modeling, that we call *relational modeling,* the features of the target phenomena are data and observation dependent, while the principle idea is that if we increase the amount of observations, these data dependent features will emerge to a set of invariant characteristics of the real phenomena.

As we mentioned before, new technologies and concepts such as ubiquitous and pervasive computing have opened up a new landscape for empirical observations and nowadays, in many application areas the conditions for relational models hold as we have an emergent network of connected instances that can be used for relational representation of the object of inquiry. From mathematical point of view this mode of representation fits very well with the concepts of probabilistic graphical models such as Markov networks and Bayesian networks that perform extremely well in coexistence with huge amount of data.

Early tries toward this direction can be referred to works of Markov (1906) in modeling natural language through sequences of written text. He claimed that only by having enough observed sequence of symbols (e.g. words), one can create a probabilistic network of relationality from each concrete instance (i.e. each symbol) to all the other instances and then, this relational network can be used as an implicit representation of that language. While in a traditional approach each instance is referred to an ideal model of its corresponding language including semantic and syntax, in a probabilistic model of the language, instances are implicitly represented by their relations to the other instances. As a result, if two words have the same function in that language, they will have similar relations with the other words. However, as later Shannon in 1948 mentioned (Shannon 1948), even after almost 40 years, the proposed relational modeling framework of Markov was not feasible as it demands for quiet large amount of observations and relatively a large computational power to process the data.

Nevertheless, the rapid growth of computational power and availability of huge amount of distributed data streams during the last decade changed the situation dramatically. The first real application of Markov networks in a large-scale problem led to the initiation of Google search engine (Brin and Page 1998) from 2000. Further, recently new applications of neural probabilistic models of the language are becoming popular, while the classical approaches in natural language processing are in catch up (Benjio *et. al* 2006 and Halevy *et. al* 2009).

With the same methodology, it is also possible to model similar problems in urban domain. For example, a Markov chain, constructed upon available GPS tracks of cars, can be used for modeling the dynamics of traffic in an urban street network (Moosavi and Hovestadt 2013), while referring to state of the art in traffic simulation methods, such as agent based models, there are lots of difficulties in tuning the behavior of artificial agents to the observed data set.

Following the same argumentation for the issue of representation in complex systems, there is another powerful data-driven modeling method, called Self Organizing Map (SOM) (Kohonnen 1982).

The main interesting property of SOM is its unique disposition for *structural learning*. Figure 4 shows the main difference of SOM to a classical way of relation (function) modeling. In simple terms, here the primary goal is to find the relation between two dimensions, based on a set of observations. In a classical way of (rational) modeling, one needs to fit a curve (a fixed structure) to a data set, while minimizing the deviations (error) from the selected curve. In other words, the curve represents the logic that integrates the observed data into a continuous relation. SOM on the other hand assumes that the logics (the argument which integrates cases) can be extracted from within the observed data – and it conserves all the logics (arguments) according to which it clusters the cases. What is optimized, in such modeling, is not how the data fits to logic, but the logic, which can integrate, as much as possible from the data.

In this sense, analogically, we might say the classical approach can be considered as a democratic set up, in which there is a global structure, tuned locally by the effect of individual instances. On the other hand, SOM provides a social environment, in which each individual instance is not reduced, but is kept active in its own individuality, while individuals can be unified into local clusters, if necessary.

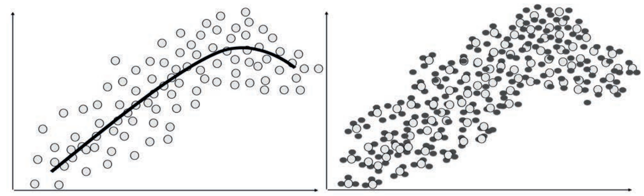

Figure 4- (left) Democratic Computing, (right) Social Computing of Self Organizing Maps (right)

This simple idea of structural learning results in a powerful generic capacity for nonlinear function approximation in coexistence with data streams (Barreto and Souza 2006). Further, it can be used as a powerful nonlinear pattern mining method.

Here again, the final model of the real phenomena is an abstraction of any potential specific model and it does not import any axiomatic or semantic specificity. Further, if the real environment is dynamic and evolving, and if we can assume the availability of dynamic data streams, then SOM is evolving along with the environment. In other words, relational models are *models in coexistence with data streams*.

These few methods are parts of a larger category of modeling approaches that together could shape new languages and new ways of encoding the real world phenomena including cities that we refer as *data literacy*.

Conclusions

Assuming that computational technologies have been the dominant factor of the last century in shaping the area of scientific urban modeling, we identified the main historical modeling capacities offered to urban modelers from computer science. Recently, as a result of ever growing ubiquitous computing systems, embedded in many daily life activities, we are observing a new emergent phenomenon, in which for the first time in the history streams of unstructured data can be seen as the new raw material of the information society. Although these data streams have inverted the process of modeling from a theory based models to data driven models, we claim that in order to be able to grasp the richness of this new capacity, we need a new kind of literacy, a technical know-how for dealing with data streams.

References

Barreto, G. A. and Souza, L. G. M. (2006). "Adaptive filtering with the self-organizing map: A performance comparison", Neural Networks, 19(6): 785-798.

Batty, M., (2005) Cities and Complexity: Understanding Cities with Cellular Automata, Agent-Based Models, and Fractals, MIT Press, Cambridge, MA & London.

Batty, M., (2008) The Size, Scale, and Shape of Cities, Science, 319, (5864, 8 February), 769 – 771

Batty, M. (2009). Urban modeling. *International Encyclopedia of Human Geography, Elsevier, Oxford*

Bellman, R. (1961). Adaptive control processes: a guided tour (Vol. 4). Princeton: Princeton university press.

Bengio, Y., Schwenk, H., Senécal, J. S., Morin, F., & Gauvain, J. L. (2006). Neural probabilistic language models. In *Innovations in Machine Learning* (pp. 137-186). Springer Berlin Heidelberg.

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 (17): 7301–7306.

Brin, S.; Page, L. (1998). "The anatomy of a large-scale hypertextual Web search engine". Computer Networks and ISDN Systems 30: 107–117.

Buhlmann, V. (2013) Die Empörung des Modells // models, outraged, Zhdk conference "David and Goliath – Models between art and architecture, November 2013 Zurich: http://wp.me/p2R9Ic-t1

Castells, M. (1989). *The informational city: information technology, economic restructuring, and the urban-regional process* (p. 15). Oxford: Blackwell.

Cranshaw, J., Schwartz, R., Hong, J. I., & Sadeh, N. M. (2012, June). The Livehoods Project: Utilizing Social Media to Understand the Dynamics of a City. In *ICWSM*.

Christaller, Walter. (1933). *Die zentralen Orte in Suddeutschland*. Jena: Gustav Fischer, 1933. (Translated (in part), by Charlisle W. Baskin, as *Central Places in Southern Germany*. Prentice Hall, 1966

Eagle, N., & Pentland, A. (2006). Reality mining: sensing complex social systems. *Personal and ubiquitous computing*, *10*(4), 255-268.

Ellerman, D. P. (1988). Category theory and concrete universals. Erkenntnis,28(3), 409-429.

Forrester, J.W.(1969). Urban dynamics. Cambridge, Mass.: MIT Press.

Foucault, M. (2002). *The order of things: An archaeology of the human sciences*. Psychology Press.

Greenfield, A. (2006). Everyware—The dawning age of ubiquitous computing.

Halevy, A., Norvig, P., & Pereira, F. (2009). The unreasonable effectiveness of data. *Intelligent Systems, IEEE*, *24*(2), 8-12.

Hansmann, U. (Ed.). (2003). *Pervasive computing: The mobile world*. Springer.

Harris, B. (1964) A Model of Locational Equilibrium for the Retail Trade' (Penn-Jersey Transportation Study, Philadelphia).

Hillier, B. and J. Hanson (1984) The Social Logic of Space, Cambridge University Press.

Hovestadt, L., Buhlmann, V., (2013), Eigen architecture, Ambra; Bilingual edition (January 2014)

Jing Yuan, et al. (2010) T-Drive: Driving Directions Based on Taxi Trajectories. In ACM SIGSPATIAL GIS 2010.

Kohonen, Teuvo (1982). "Self-Organized Formation of Topologically Correct Feature Maps". Biological Cybernetics 43 (1): 59–69.

Lee D.B. (1973), "A Requiem for Large Scale Modeling", Journal of The American Institute of Planners, 39 (1973) (3), pp. 163–178.

Lowry, I S. (1964) A Model of Metropolis RAND Memorandum 4025-RC.

Lynch, K. (1984). *Good city form*. MIT press.

Markov, A.A. (1906) ‚Rasprostranenie zakona bol'shih chisel na velichiny, zavisyaschie drug ot druga, Izvestiya Fiziko-matematicheskogo obschestva pri Kazanskom universitete, 2-ya seriya 15 (94) 135–156.

Moosavi, V., & Hovestadt, L. (2013). Modeling urban traffic dynamics in coexistence with urban data streams. In Proceedings of the 2nd ACM SIGKDD International Workshop on Urban Computing (p. 10). ACM.

Moosavi, Vahid. 2014, "Beyond Rational Modeling", European Meetings on Cybernetics and Systems Research, Vienna, 2014, http://emcsr.net/book-of-abstracts/.

Noulas, A., Scellato, S., Lambiotte, R., Pontil, M., & Mascolo, C. (2012). A tale of many cities: universal patterns in human urban mobility. *PloS one*, *7*(5), e37027.

Paulos, Eric and R.J. Honicky, and Ben Hooker,(2008) Citizen Science: Enabling Participatory Urbanism in Handbook of Research on Urban Informatics: The Practice and Promise of the Real-Time City. Edited by Marcus Foth Hershey, PA: Information Science Reference, IGI Global.

Shane, D. G. (2005). *Recombinant urbanism: conceptual modeling in architecture, urban design, and city theory*. Chichester: Wiley.

Shannon, Claude E. (1948). "A Mathematical Theory of Communication". Bell System Technical Journal 27 (3): 379–423.

Simini, F., González, M. C., Maritan, A., & Barabási, A. L. (2012). A universal model for mobility and migration patterns. *Nature*, *484*(7392), 96-100.

Tobler, W. (1979) Cellular geography, Philosophy in Geography. Ed. Gale, S and G Olsson. D. Reidel, Boston. 379-386.

Von Thünen, J. H. (1966). *Isolated state: an English edition of Der isolierte Staat*. Pergamon Press.

Waddell, P. and G. Ulfarsson.(2004).Introduction to urban simulation: design and development of operational models, Handbook of Transport Geography and Spatial Systems, Volume 5 (Handbooks in Transport). Ed. Hensher, D, K Button, K Haynes and P Stopher. Elsevier Science, 203-236.

Weisberg, M. (2007). Three kinds of idealization. *The Journal of Philosophy*, 639-659.

Wilson, A. (2012). *Urban Modelling: Critical Concepts in Urban Studies*. Routledge.

Wolman, A. (1965). The Metabolism of the City. Scientific American. 213, 179-190.