Methods and Measures for Analyzing Complex Urban Street Networks

*ROUGH DRAFT IN PROGRESS*

Geoff Boeing

Department of City and Regional Planning

University of California, Berkeley

November 2016

Table of Contents

[Table of Contents 2](#_Toc466395498)

[List of Figures 5](#_Toc466395499)

[List of Tables 8](#_Toc466395500)

[Acknowledgements 9](#_Toc466395501)

[1. Introduction 10](#_Toc466395502)

[1.1. Abstract 11](#_Toc466395503)

[1.2. Motivation 11](#_Toc466395504)

[1.3. Context of the Study 11](#_Toc466395505)

[1.4. Organization and Contribution by Chapter 11](#_Toc466395506)

[1.4.1. Chapter 1: Introduction 12](#_Toc466395507)

[1.4.2. Chapter 2: Foundations of the Nonlinear Paradigm 12](#_Toc466395508)

[1.4.3. Chapter 3: Complexity and Cities 13](#_Toc466395509)

[1.4.4. Chapter 4: Methods for Measuring the Complexity Outcomes of Urban Design 14](#_Toc466395510)

[1.4.5. Chapter 5: Acquiring, Constructing, Analyzing, and Visualizing Street Networks 16](#_Toc466395511)

[1.4.6. Chapter 6: Case Study: Portland 17](#_Toc466395512)

[1.4.7. Chapter 6: Multi-scale Analysis of Urban Street Networks 18](#_Toc466395513)

[1.4.8. Chapter 7: Conclusion 18](#_Toc466395514)

[2. Foundations of the Nonlinear Paradigm 19](#_Toc466395515)

[2.1. Abstract 20](#_Toc466395516)

[2.2. Introduction 20](#_Toc466395517)

[2.3. Background and Model 24](#_Toc466395518)

[2.4. System Bifurcations 29](#_Toc466395519)

[2.5. Fractals and Strange Attractors 33](#_Toc466395520)

[2.6. Chaos and Randomness 41](#_Toc466395521)

[2.7. The Butterfly Effect 44](#_Toc466395522)

[2.8. Discussion 47](#_Toc466395523)

[2.9. Conclusion 51](#_Toc466395524)

[3. Complexity and Cities 53](#_Toc466395525)

[3.1. Abstract 54](#_Toc466395526)

[3.2. Introduction 54](#_Toc466395527)

[3.3. Systems and Dynamics 57](#_Toc466395528)

[3.4. Measures of Complexity 60](#_Toc466395529)

[3.5. Equilibrium and Stability 62](#_Toc466395530)

[3.6. Emergence, Self-Organization, and Resilience 67](#_Toc466395531)

[3.7. Networks 71](#_Toc466395532)

[3.8. Conclusion 75](#_Toc466395533)

[4. Methods for Measuring the Complexity Outcomes of Urban Design 76](#_Toc466395534)

[4.1. Abstract 77](#_Toc466395535)

[4.2. Introduction 77](#_Toc466395536)

[4.3. Background: Complexity in Urban Design 79](#_Toc466395537)

[4.3.1. Urban Design, Livability, and Complexity 80](#_Toc466395538)

[4.3.2. The Neighborhood Scale 82](#_Toc466395539)

[4.3.3. Designing for Complexity 84](#_Toc466395540)

[4.4. Measures of Complexity in Urban Form and Design 86](#_Toc466395541)

[4.4.1. Overview 86](#_Toc466395542)

[4.4.2. Temporal Measures of Urban Form 89](#_Toc466395543)

[4.4.3. Visual Complexity of Urban Form 91](#_Toc466395544)

[4.4.4. Spatial Measures of Urban Form 94](#_Toc466395545)

[4.4.5. Structural Measures of Urban Form: Fractal 97](#_Toc466395546)

[4.4.6. Structural Measures of Urban Form: Network 100](#_Toc466395547)

[4.5. Preliminary Typology of Complexity Measures 106](#_Toc466395548)

[4.6. Discussion 107](#_Toc466395549)

[5. Acquiring, Constructing, Analyzing, and Visualizing Street Networks 110](#_Toc466395550)

[5.1. Abstract 111](#_Toc466395551)

[5.2. Introduction 111](#_Toc466395552)

[5.3. Background 113](#_Toc466395553)

[5.3.1. Street network analysis 113](#_Toc466395554)

[5.3.2. Representation of street networks 113](#_Toc466395555)

[5.3.3. Current tool landscape 116](#_Toc466395556)

[5.3.4. Summary 119](#_Toc466395557)

[5.4. OSMnx: Functionality and comparison to existing tools 120](#_Toc466395558)

[5.4.1. Acquiring administrative place boundaries 121](#_Toc466395559)

[5.4.2. Download and construct street networks 122](#_Toc466395560)

[5.4.3. Correct and simplify network topology 127](#_Toc466395561)

[5.4.4. Save street networks to disk 130](#_Toc466395562)

[5.4.5. Analyze street networks 131](#_Toc466395563)

[5.4.6. Summary 137](#_Toc466395564)

[5.5. Discussion 137](#_Toc466395565)

[6. Case Study: Portland, Oregon 139](#_Toc466395566)

[6.1. Abstract 140](#_Toc466395567)

[6.2. Introduction 140](#_Toc466395568)

[6.3. Methods 141](#_Toc466395569)

[6.4. Findings 142](#_Toc466395570)

[6.5. Discussion 148](#_Toc466395571)

[7. Multi-scale analysis of urban street networks 150](#_Toc466395572)

[7.1. Abstract 151](#_Toc466395573)

[7.2. Introduction 151](#_Toc466395574)

[7.3. Methods 153](#_Toc466395575)

[7.4. Analysis of Metropolitan-Scale Street Networks 155](#_Toc466395576)

[7.5. Analysis of Municipal-Scale Street Networks 155](#_Toc466395577)

[7.6. Analysis of Neighborhood-Scale Street Networks 155](#_Toc466395578)

[7.7. Intrametropolitan Neighborhood-Scale Analysis of San Francisco 155](#_Toc466395579)

[7.8. Discussion 155](#_Toc466395580)

[8. Conclusion 156](#_Toc466395581)

[8.1. Summary of Key Findings 157](#_Toc466395582)

[8.2. Contributions 157](#_Toc466395583)

[8.2.1. Contribution to the Literature 157](#_Toc466395584)

[8.2.2. Contribution to Planning Practice 157](#_Toc466395585)

[8.3. Future Research 157](#_Toc466395586)

[Bibliography 159](#_Toc466395587)

List of Figures

[Figure 2.1. Time series graph of the logistic map with 7 growth rate parameter values over 20 generations. 28](#_Toc466395822)

[Figure 2.2. Bifurcation diagram of 100 generations of the logistic map for 1,000 growth rate parameter values between 0 and 4. The vertical slice above each growth rate depicts the system’s attractor at that rate. 30](#_Toc466395823)

[Figure 2.3. Bifurcation diagram of 100 generations of the logistic map for 1,000 growth rate parameter values between 2.8 and 4. The vertical slice above each growth rate depicts the system’s attractor at that rate. 31](#_Toc466395824)

[Figure 2.4. Bifurcation diagram of 100 generations of the logistic map for 1,000 growth rate parameter values between 3.7 and 3.9. The system moves from order to chaos and back again as the growth rate is adjusted. 32](#_Toc466395825)

[Figure 2.5. Bifurcation diagram of 100 generations of the logistic map for 1,000 growth rate parameter values between 3.84 and 3.856. This is the same structure that we saw earlier at the macro-level in Figure 2.3, because chaotic systems’ strange attractors are fractal. 34](#_Toc466395826)

[Figure 2.6. Phase diagrams of the logistic map over 200 generations for growth rate parameter values of 2.9 (A), 3.5 (B), 3.56 (C), and 3.57 (D). When the parameter is set to 2.9, the model converges at a single fixed-point. When the parameter is set to 3.5 or higher, the model oscillates over four points, then eight, and on and on as it bifurcates. 35](#_Toc466395827)

[Figure 2.7. Cropped phase diagrams of the logistic map over 200 generations for (A) a growth rate parameter value of 3.9 and (B) 50 growth rate parameter values between 3.6 and 4 (the chaotic regime), each with its own colored line. 37](#_Toc466395828)

[Figure 2.8. Cobweb plots of the logistic map for growth rate parameter values of (A) 1, (B) 2.7, (C) 3.5, (D) 3.9. The diagonal gray identity line represents y=x, the red curve represents the logistic map as y=f(x) for each of the four parameter values, and the blue cobweb line represents the system’s trajectory over 100 generations. 40](#_Toc466395829)

[Figure 2.9. Plot of two time series, one deterministic/chaotic from the logistic map (blue), and one random (red). 42](#_Toc466395830)

[Figure 2.10. Phase diagrams of the time series in Figure 2.9. 10B is a three-dimensional state space version of the two-dimensional 10A. 43](#_Toc466395831)

[Figure 2.11. Two different viewing perspectives of a single three-dimensional phase diagram of the logistic map over 200 generations for 50 growth rate parameter values between 3.6 and 4, each with its own colored line. 43](#_Toc466395832)

[Figure 2.12. Cobweb plots of the logistic map pulling initial population values of 0.1 (A), 0.5 (B), and 0.9 (C) into the same fixed-point attractor over time. At this growth rate parameter value of 2.7, the Lyapunov is negative. 45](#_Toc466395833)

[Figure 2.13. Plot of two time series with identical dynamics, one starting at an initial population value of 0.5 (blue) and the other starting at 0.50001 (red). At this growth rate parameter value of 3.9, the Lyapunov is positive – thus the system is chaotic and we can see the nearby points diverge over time. 46](#_Toc466395834)

[Figure 3.1 An example phase space diagram of the evolution of the Lorenz system over time. 59](#_Toc466395835)

[Figure 5.1 Administrative boundary vector geometries retrieved for A) Berkeley, California, and B) Zambia, Zimbabwe, and Botswana. 122](#_Toc466395836)

[Figure 5.2 Street networks created by A) bounding box, B) address and network distance, and C) neighborhood polygon 124](#_Toc466395837)

[Figure 5.3 The drivable street network for municipal Los Angeles, created by simply passing the query phrase "Los Angeles, CA, USA" into OSMnx. 126](#_Toc466395838)

[Figure 5.4 Street networks for A) Modena, Italy, B) Belgrade, Serbia, C) central Maputo, Mozambique, and D) central Tunis, Tunisia. 127](#_Toc466395839)

[Figure 5.5 A) the original graph, B) non-graph-theoretic nodes highlighted in red and true intersections in blue, C) strictly simplified network, with self-loops noted in magenta, D) non-strictly simplified network. 130](#_Toc466395840)

[Figure 5.6 Street network for metropolitan New York from OSMnx saved and loaded in QGIS as an ESRI shapefile (above) and in Adobe Illustrator as SVG (below). 131](#_Toc466395841)

[Figure 5.7 OSMnx calculates the shortest network path between two points, accounting for one-way routes, and plots it. 135](#_Toc466395842)

[Figure 5.8 OSMnx visualizes the spatial distribution of cul-de-sacs in Piedmont, California. 136](#_Toc466395843)

[Figure 5.9 One square mile of each city, created and plotted automatically by OSMnx in the style of Allan Jacobs (1995). 136](#_Toc466395844)

[Figure 6.1 Three 0.5 km2 sections (each centered on the coordinates presented in section 6.3) of the street network in Portland, Oregon projected and plotted by OSMnx. 143](#_Toc466395845)

[Figure 6.4 Distribution of intersection types in each street network section: x-axis is number of streets and y-axis is the proportion of intersections with that number. 146](#_Toc466395846)

[Figure 6.5 Three 0.5 km2 sections of the street network in Portland, Oregon. Nodes colored by betweenness centrality from lowest (dark) to highest (light). 147](#_Toc466395847)

List of Tables

[Table 2.1. Population values produced by the logistic map with 7 growth rate parameter values over 20 generations. 27](#_Toc466395677)

[Table 4.1 Typology of measures of the complexity of urban form/design. 107](#_Toc466395678)

[Table 5.1 Network metrics and statistics automatically calculated by OSMnx. 134](#_Toc466395679)

[Table 6.1 Descriptive statistics for three street network sections in the city of Portland, Oregon. 145](#_Toc466395680)

Acknowledgements

# Introduction

## Abstract

This chapter introduces the motivation for and context of the study presented in this dissertation. It also summarizes the organization and contribution of each of the subsequent chapters in the dissertation.

## Motivation

## Context of the Study

## Organization and Contribution by Chapter

This dissertation begins and ends with introductory and concluding chapters that book-end its six central substantive chapters. These six chapters unpack the foundations of the nonlinear paradigm, contextualize urban street network analysis within theories of complexity, create a typology for measuring the complexity of urban design, present a new method for acquiring, constructing, analyzing and visualizing street networks, and conduct a multi-scale analysis of urban street networks across the United States.

### Chapter 1: Introduction

This chapter has thus far introduced the motivation for and context of the study presented in this dissertation. The remainder of this chapter summarizes the organization and contribution of each of the subsequent chapters in this dissertation.

### Chapter 2: Foundations of the Nonlinear Paradigm

Chapter 2 serves provides a background for the rest of the dissertation and has two primary aims. First it lays the foundation of the complexity theories of cities presented in chapter 3 by introducing the fundamentals of nonlinear dynamics, chaos, fractals, self-similarity, and the limits of prediction. It does so interdisciplinarily through several visualization methods to analyze and understand system behavior. Second it presents Pynamical, a new tool to visualize and explore nonlinear dynamical systems’ behavior. Nearly all nontrivial real-world systems are nonlinear dynamical systems. The modern study of complex systems evolved from initial explorations of the surprising behavior of such systems. Although the social sciences are increasingly studying society and cities through the lens of these types of systems, seminal concepts remain murky or loosely adopted in the literature. In particular, this chapter introduces systems, dynamics, self-similarity, and prediction to set up the discussion in chapter 3 of complexity, cities, and the study of networks.

This chapter makes two contributions: one theoretical, one methodological. First, it reviews and theory of nonlinearity for the qualitative analysis of nonlinear dynamical systems’ behavior to an interdisciplinary body of urban scholars and planners. Most formal treatments of chaos and nonlinear dynamics in the scholarly literature are densely technical and geared toward an exclusive audience of mathematicians and physicists. For this chapter, rather, readers require only a familiarity with algebra. Second, this chapter makes a methodological contribution by presenting Pynamical, a new tool to visualize and explore nonlinear dynamical systems’ behavior. Comparable tools usually must be developed from scratch or rely on expensive commercial software such as MATLAB. Developing tools for exploring, understanding, and visualizing dynamical systems in Python makes them available to a much wider audience of systems analysts, researchers, and students. Pynamical provides a fast, simple, reusable, extensible, free, and open-source new means for exploring system behavior – particularly for the qualitative analysis of such systems in research and pedagogy.

### Chapter 3: Complexity and Cities

Building on the background of chapter 2, chapter 3 presents the theoretical framework of complex systems and cities, culminating in network theory and analysis – the primary lens this study uses in all subsequent chapters. Discussions of complexity and complex systems have appeared throughout the planning literature for years. These principles have been applied everywhere from the communicative turn and collaborative rationality, to cellular automata and agent-based urban models, to the design of resilient, livable neighborhoods. However, the interdisciplinary appeal and trendiness of complexity in the social sciences has resulted in a bit of a morass of ambiguous terminology, internal inconsistencies, and overloaded concepts open to multiple interpretations.

Unlike all the other substantive chapters in this dissertation, this chapter makes neither an empirical nor a methodological contribution. However, it offers a theoretical contribution to the urban planning literature by unpacking the key foundational concepts of complex systems and network science in a brief, straightforward manner. It provides explanatory examples of these concepts familiar to scholars and practitioners not already versed in the technical science of complexity. Most relevant to this present study, this chapter presents the theory of networks and network analysis that form the foundation of the remaining chapters.

### Chapter 4: Methods for Measuring the Complexity Outcomes of Urban Design

Building on the theories of complexity and networks presented in chapter 3, chapter 4 develops a typology of measures for assessing the complexity of the urban built form. In particular, it extends quantitative methods from network science, ecosystems studies, fractal geometry, and information theory to the practice of neighborhood-scale urban design and the analysis of its qualitative human experience. Metrics at multiple scales are scattered throughout these bodies of literature and have useful applications in analyzing the built form that results from local planning and design processes. Rich linkages between complexity theory and urban design have been underexplored by researchers at the neighborhood and street scales – the scales of daily human experience. The urban design literature frequently cites the value of *complexity* in neighborhood design, but these arguments often lack the formalism found in complex systems science. If neighborhood complexity is considered important, how might we interpret it and how might it be assessed?

This chapter unpacks the connections between neighborhood-scale built form and measures of its complexity. It contributes a new typology of tools and metrics from different scientific disciplines to assess measures of complexity that apply to urban form and particularly at urban design’s scale of intervention. In particular, the measures of network structure characterize the complexity of the circulation network in terms of density, permeability, resilience, and connectedness. These attributes influence the way an urban system’s physical links can structure complex interactions, connections, and dynamics. The analytical framework developed here is generalizable to empirical research of multiple neighborhood types and design standards. In particular, network-analytic measures in this typology are operationalized in the next chapter, and applied empirically in the subsequent two empirical chapters.

### Chapter 5: Acquiring, Constructing, Analyzing, and Visualizing Street Networks

Scholars have studied street networks in various ways. However, there are some limitations to the current urban planning/street network analysis literature. To address these challenges, this study presents a new tool to make the collection of data and creation and analysis of street networks simple, consistent, and automatable. OSMnx is a new Python package that downloads administrative boundary shapes and street networks from OpenStreetMap. OSMnx contributes five significant new capabilities for researchers and city planners: first, the automatic downloading of administrative place boundaries and shapefiles; second, the tailored and automated downloading and constructing of street networks from OpenStreetMap; third, the automated correction and simplification of network topology; fourth, the ability to save street networks to disk as shapefiles, GraphML, or SVG files; and fifth, the ability to analyze street networks, calculate routes, project and visualize the networks, and calculate network metrics and statistics. These metrics and statistics include both those common in urban design and transportation studies, and metrics that measure the structure and topology of the network.

This chapter makes two primary methodological contributions. First, it presents new methods and tools for acquiring, constructing, correcting, projecting, analyzing, and visualizing street networks. Second, it adapts measures from traditional network analysis to make them better-suited to accurately describing the physical form of street intersections and network connectivity.

### Chapter 6: Case Study: Portland

This short chapter presents a small case study to simply but plainly demonstrate the usage of OSMnx for research. It collects three small half-kilometer sections of the street network in different neighborhoods in Portland, Oregon to perform a cross-sectional analysis. This scale of analysis and sample size are small, but they provide simple, comprehensible examples to illustrate the network concepts presented in chapter 3, the network measures presented in chapter 4, and the methodological tool presented in chapter 5. This chapter thus serves to tie these preceding threads together empirically.

This chapter has two aims. First, as discussed, it demonstrates the functionality of OSMnx with a simple case study. Second, it presents empirical findings of three street network sections in Portland, Oregon and uses the quantitative measures to compare and contrast these network sections. It introduces these neighborhoods first from a brief qualitative and historical perspective, then explores their comparative quantitative measures of network complexity and structure. Then it discusses these empirical findings and what insights may be drawn from them. This discussion offers some insights and suggestions arising from these findings. It identifies significant chokepoints in the suburban network and demonstrates how there could be substantial gains in network robustness if one-way streets in the dense, orthogonal downtown were converted to two-way streets. This small case study serves to demonstrate OSMnx simply before embarking on the large multi-scale analysis of 27,000 street networks in the following chapter.

### Chapter 7: Multi-scale Analysis of Urban Street Networks

### Chapter 8: Conclusion

The dissertation concludes with a summary of key findings, a discussion of their contribution to the literature, and trajectories for future research.

# Foundations of the Nonlinear Paradigm

## Abstract

Nearly all nontrivial real-world systems are nonlinear dynamical systems. Chaos describes certain nonlinear dynamical systems that have a very sensitive dependence on initial conditions. Chaotic systems are always deterministic and may be very simple, yet produce completely unpredictable and divergent behavior. The modern study of complex systems evolved from these initial explorations, and although the social sciences are increasingly studying these types of systems, seminal concepts remain murky or loosely adopted. This chapter has two primary aims. First it introduces the foundations of nonlinear dynamics, chaos, fractals, self-similarity, and the limits of prediction through several visualization methods to analyze and understand system behavior. Second it presents Pynamical, a new tool to visualize and explore nonlinear dynamical systems’ behavior.

## Introduction

The modern study of complex systems evolved from initial explorations of nonlinear dynamical systems in the second half of the twentieth century, in the then-nascent field of chaos theory. Chaos theory is a branch of mathematics that deals with nonlinear dynamical systems. A *system* is simply a set of interacting components that form a larger whole. *Nonlinear* means that due to feedback or multiplicative effects between the components, the whole becomes something greater than the mere sum of its individual parts. Lastly, *dynamical* means the system changes over time based on its current state. Nearly every nontrivial real-world system is a nonlinear dynamical system. Chaotic systems are a type of nonlinear dynamical systems that may contain very few interacting parts and may follow simple rules, but all have a very sensitive dependence on their initial conditions (Hastings, Hom, Ellner, Turchin, & Godfray, 1993; Rickles, Hawe, & Shiell, 2007). One might expect that any simple deterministic system would produce easily predictable behavior. Yet despite their deterministic simplicity, over time these systems can produce wildly unpredictable, divergent, and fractal (i.e., infinitely detailed and self-similar without ever actually repeating) behavior due to that sensitivity. Forecasting such systems’ futures thus requires an impossible precision of measurement and computation. Chaos fundamentally indicates that there are limits to knowledge and prediction because some futures may be unknowable with any precision. Further, interventions into a system may have unpredictable outcomes even if the intervention is very minor, as tiny effects can compound (or be damped) nonlinearly over time.

Real-world chaotic and fractal systems span the spectrum from leaky faucets (Suetani, Soejima, Matsuoka, Parlitz, & Hata, 2012), to ferns (Singh, Mishra, & Sinkala, 2012), to heart rates (Babbs, 2014; Glass, 2009; Hoshi, Pastre, Vanderlei, & Godoy, 2013), to cryptography (Hong & Dong, 2010; Makris & Antoniou, 2012). Many scholars have studied the implications of nonlinearity, chaos, and fractals for the social sciences, including sociology (Guastello, 2013; Richards, 1996), urban studies (Batty & Longley, 1994; Batty & Xie, 1999; Benguigui, Czamanski, Marinov, & Portugali, 2000; Y. Chen & Zhou, 2008; Shen, 2002), economics (W.-C. Chen, 2008; Guégan, 2009; Oxley & George, 2007; Puu, 2013; Rosser, Jr., 1996), architecture (Hamouche, 2009; Ostwald, 2013), and city planning (Batty, 2013; Batty & Marshall, 2012; Cartwright, 1991; Innes & Booher, 2010). One constant throughout the interdisciplinary history of nonlinear dynamical systems study is that nonlinear systems are extremely difficult to solve analytically because they cannot be broken down into constituent parts, solved individually, then recombined as a solution. Scientists have instead relied heavily on visual and qualitative approaches – a perspective first developed by Henri Poincaré in the late 1800s – to discover and analyze the fascinating dynamics of nonlinearity (Alpigini, 2004; Layek, 2015). Information visualization helps analysts detect and examine hidden structure in complex data sets (C. Chen, 2006). In particular, few fields have drawn as heavily from visualization as nonlinear dynamics and chaos have for their pivotal discoveries, from Lorenz’s first visualization of strange attractors (Lorenz, 1963), to May’s groundbreaking bifurcation diagrams (May, 1976), to phase diagrams for discerning higher-dimensional hidden structures in data (Packard, Crutchfield, Farmer, & Shaw, 1980). Such nonlinear analysis is particularly useful yet underutilized for exploring time series (Bradley, 2003; Bradley & Kantz, 2015). These methods in turn have broad applicability to visual information analysis and the interdisciplinary study of nonlinear and complex systems.

This chapter introduces nonlinearity through the methods of data visualization, using a logistic model to dissect the terminology, visualize pertinent features of chaos and fractals, and discuss wide-ranging implications for knowledge and prediction. It has two primary aims. First, it introduces the foundations of nonlinear dynamics, chaos, fractals, self-similarity, and the limits of prediction. Although the social sciences are increasingly studying these types of systems, some of the seminal concepts remain murky or loosely adopted in the theoretical literature (Chettiparamb, 2006). Most *formal* treatments of chaos and nonlinear dynamics in the scholarly literature are densely technical and geared toward an exclusive audience of mathematicians and natural scientists. For this chapter, rather, readers require only a familiarity with algebra. We thus do not cover the rigorous mathematical underpinnings of chaos and nonlinear dynamics, but the references throughout cite both the original foundational publications in this field as well as recent scholarly developments. Interested readers will be well-rewarded in consulting these works. Second, this chapter presents Pynamical, a new tool to visualize and explore nonlinear dynamical systems’ behavior. Comparable tools usually must be developed from scratch or rely on expensive commercial software such as MATLAB (Tomida, 2008). Pynamical provides a fast, simple, reusable, extensible, free, and open-source new means for exploring system behavior – particularly for the qualitative analysis of such systems in research and pedagogy.

The following section provides a background to the logistic map and the concepts of system dynamics and attractors. Then we introduce several information visualization techniques to explore qualitative system behavior, bifurcations, the path to chaos, fractals, and strange attractors. We investigate the difference between chaos and randomness. Finally, we visualize the famous butterfly effect and conclude with a discussion of its implications for scientific prediction and complexity. All of these models and visualizations are developed in Python using Pynamical; for readability, we reserve the technical details of its functionality for the appendix.

## Background and Model

Edward Lorenz, the father of chaos theory (Stewart, 2000), once described chaos as “when the present determines the future, but the approximate present does not approximately determine the future” (Danforth, 2013). Lorenz first discovered chaos by accident while developing a simple mathematical model of atmospheric convection, using three ordinary differential equations (Lorenz, 1963). He found that nearly indistinguishable initial conditions could produce completely divergent outcomes, rendering weather prediction impossible beyond a time horizon of about a fortnight (Gleick, 1991).

How can this possibly happen with a simple deterministic system? We will explore an example using the *logistic map*, a model based on the common s-curve logistic function that shows how a population grows slowly, then rapidly, before tapering off as it reaches its environment’s carrying capacity (W. Li, Wang, & Su, 2011; May, 1974). The logistic function uses a differential equation that treats time as continuous. The logistic map instead uses a difference equation to look at discrete time steps (Pastijn, 2006; Strogatz, 2014). It is called the logistic *map* because it maps the population value at any time step to its value at the next time step: *xt*+1 = *r xt* (1–*xt*). This nonlinear equation defines the rules, or *dynamics*, of our system: *x* represents the population at some time *t*, and *r* represents the growth rate. Thus the population level at any given time is a function of the growth rate parameter and the previous time step’s population level. If the growth rate is set too low, the population will die out and go extinct. Higher growth rates might settle toward a stable value or fluctuate across a series of population booms and busts.

Chaos can manifest itself in both continuous (i.e., with dynamics defined by differential equations) and discrete (i.e., with dynamics defined by an iterated map) nonlinear dynamical systems. The logistic map is a simple, one-dimensional, discrete equation that produces chaos at certain growth rates. We will explore this in depth momentarily, but first, we use Pynamical to run the logistic model for 20 time steps (we will henceforth call these recursive iterations of the equation *generations*) for growth rate parameter values of 0.5, 1, 1.5, 2, 2.5, 3, and 3.5. Table 1 presents the results. The columns represent growth rates and the rows represent generations. The model always starts with a population level of 0.5 and represents population as a ratio between 0 (extinction) and 1 (the maximum carrying capacity of our system). If we trace down the column in Table 1 under growth rate 1.5, we see that the population level eventually settles toward a final value of 0.333 after several generations. In the column for growth rate 2, we see an unchanging population level of 0.5 across every generation. This makes sense in the real world – if two parents produce two children, the overall population will neither grow nor shrink. Thus a growth rate parameter value of 2 represents the replacement rate.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Generation | r = 0.5 | r = 1.0 | r = 1.5 | r = 2.0 | r = 2.5 | r = 3.0 | r = 3.5 |
| 1 | 0.500 | 0.500 | 0.500 | 0.500 | 0.500 | 0.500 | 0.500 |
| 2 | 0.125 | 0.250 | 0.375 | 0.500 | 0.625 | 0.750 | 0.875 |
| 3 | 0.055 | 0.188 | 0.352 | 0.500 | 0.586 | 0.562 | 0.383 |
| 4 | 0.026 | 0.152 | 0.342 | 0.500 | 0.607 | 0.738 | 0.827 |
| 5 | 0.013 | 0.129 | 0.338 | 0.500 | 0.597 | 0.580 | 0.501 |
| 6 | 0.006 | 0.112 | 0.335 | 0.500 | 0.602 | 0.731 | 0.875 |
| 7 | 0.003 | 0.100 | 0.334 | 0.500 | 0.599 | 0.590 | 0.383 |
| 8 | 0.002 | 0.090 | 0.334 | 0.500 | 0.600 | 0.726 | 0.827 |
| 9 | 0.001 | 0.082 | 0.334 | 0.500 | 0.600 | 0.597 | 0.501 |
| 10 | 0.000 | 0.075 | 0.333 | 0.500 | 0.600 | 0.722 | 0.875 |
| 11 | 0.000 | 0.069 | 0.333 | 0.500 | 0.600 | 0.603 | 0.383 |
| 12 | 0.000 | 0.065 | 0.333 | 0.500 | 0.600 | 0.718 | 0.827 |
| 13 | 0.000 | 0.060 | 0.333 | 0.500 | 0.600 | 0.607 | 0.501 |
| 14 | 0.000 | 0.057 | 0.333 | 0.500 | 0.600 | 0.716 | 0.875 |
| 15 | 0.000 | 0.054 | 0.333 | 0.500 | 0.600 | 0.610 | 0.383 |
| 16 | 0.000 | 0.051 | 0.333 | 0.500 | 0.600 | 0.713 | 0.827 |
| 17 | 0.000 | 0.048 | 0.333 | 0.500 | 0.600 | 0.613 | 0.501 |
| 18 | 0.000 | 0.046 | 0.333 | 0.500 | 0.600 | 0.711 | 0.875 |
| 19 | 0.000 | 0.044 | 0.333 | 0.500 | 0.600 | 0.616 | 0.383 |
| 20 | 0.000 | 0.042 | 0.333 | 0.500 | 0.600 | 0.710 | 0.827 |

Table 2.1. Population values produced by the logistic map with 7 growth rate parameter values over 20 generations.

Figure 2.1 visualizes the resulting time series as a graph produced by Pynamical, with time on the *x*-axis and the system state on the *y*-axis. This graph visualizes how the population changes over time at different growth rates. For instance, the violet line for growth rate 0.5 quickly drops to zero: the population dies out. The teal line that represents a growth rate of 2 (the replacement rate) stays steady at a population level of 0.5. The growth rates of 3 and 3.5 are more interesting. While the green line for growth rate 3 seems to slowly converge toward a stable value, the yellow line for growth rate 3.5 just seems to repeatedly bounce around four different values.

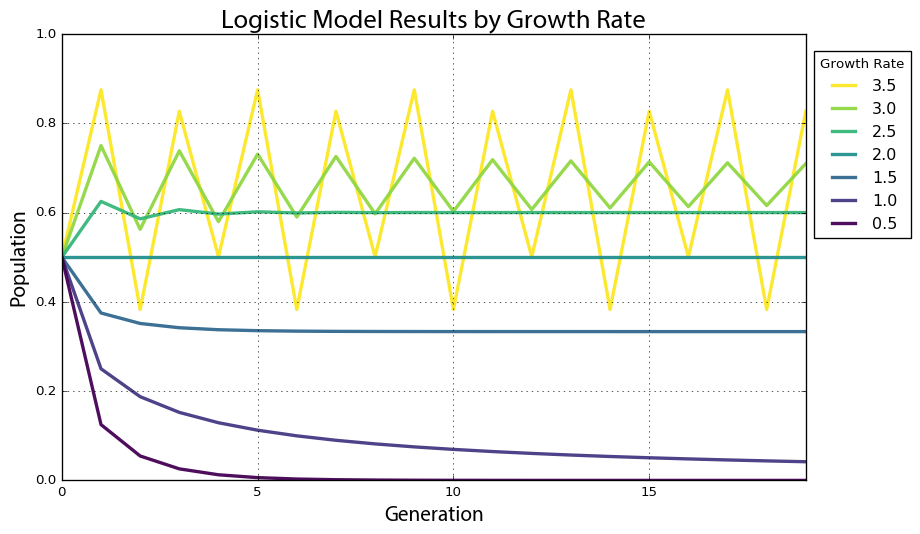


Figure 2.1. Time series graph of the logistic map with 7 growth rate parameter values over 20 generations.

An *attractor* is the value, or set of values, that a system settles toward over time. When the growth rate parameter is set to 0.5, the system has a *fixed-point attractor* at population level 0, as depicted by the violet line dropping to 0. In other words, the population value is drawn toward a stable equilibrium of 0 over time as the model iterates: the logistic equation maps the value of a fixed-point attractor to itself. When the growth rate parameter is set to 3.5, the system oscillates between four values as depicted by the yellow line. This oscillating attractor is called a *limit cycle*. But when we adjust the growth rate parameter in this model beyond 3.57, we witness the onset of chaos. A chaotic system has a *strange attractor*, around which the system oscillates forever without ever repeating itself or settling into a steady state of behavior (Ruelle & Takens, 1971; Shilnikov, 2002). It never produces the same value twice and its structure is fractal, meaning the same patterns exist at every scale no matter how much we zoom into it (Grebogi, Ott, & Yorke, 1987).

## System Bifurcations

To show this more clearly, we run the logistic model again, this time for 200 generations across 1,000 growth rate values between 0 and 4. When we produced the plot in Figure 2.1, we had only 7 growth rates. This time we have 1,000 so we need to visualize the results in a different way to make them comprehensible, using a *bifurcation diagram* that visualizes a system’s attractors as a function of some parameter (Gershenson, 2004; May, 1976; Wu & Baleanu, 2014). The bifurcation diagram in Figure 2.2 represents 1,000 discrete vertical slices, each corresponding to one of 1,000 growth rate parameter values evenly spaced between 0 and 4. To produce each of these visual slices, Pynamical ran the model 200 times then threw away the first 100 results, leaving just the final 100 generations for each growth rate. Each vertical slice thus visualizes the population values that the logistic map settles toward over time (i.e., the attractor) for that parameter value.

In Figure 2.2 we can see that for growth rates less than 1, the system always eventually collapses to zero (extinction). For growth rates between 1 and 3, the system always settles into an exact, stable population level. For instance, in the vertical slice above growth rate 2.5, there is only one population value represented (0.6) and it corresponds precisely to where the line for growth rate 2.5 settles in Figure 2.1’s time series graph. At this parameter value, the system’s attractor is a fixed point at 0.6. But for some growth rates, such as 3.9, the plot in Figure 2.2 shows 100 different values – in other words, a different value for each of its 100 generations. Here the system never settles into a fixed point or a limit cycle.

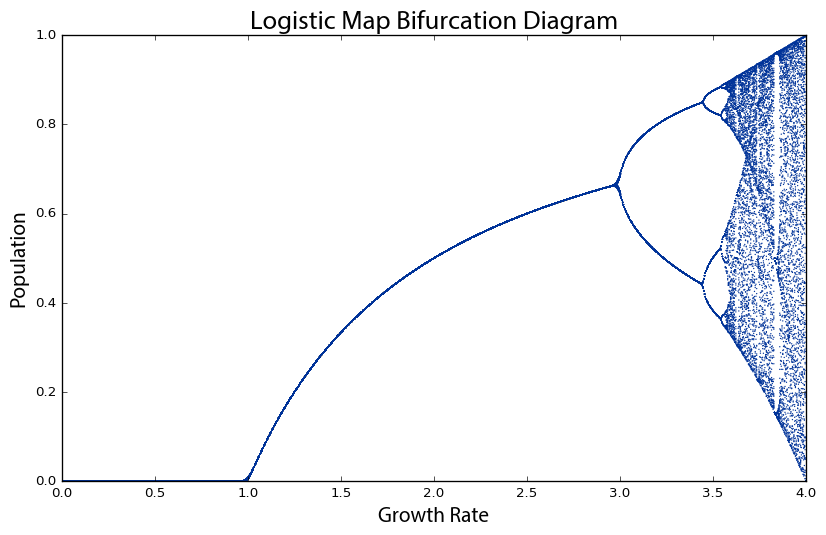


Figure 2.2. Bifurcation diagram of 100 generations of the logistic map for 1,000 growth rate parameter values between 0 and 4. The vertical slice above each growth rate depicts the system’s attractor at that rate.

Why is this visualization called a bifurcation diagram? If we zoom into the growth rates between 2.8 and 4 to see what is happening at a finer scale (Figure 2.3), the possible population values fork into two discrete paths at the vertical slice above growth rate 3. At growth rate 3.2, the system oscillates exclusively between two population values: one around 0.5 and the other around 0.8. Thus, at that growth rate, applying the logistic map to one of these two population values yields the other. Just beyond growth rate 3.4, the diagram bifurcates again into *four* paths. This corresponds to the yellow line in Figure 2.1: when the growth rate parameter is set to 3.5, the system oscillates over *four* population values. These are *periods*, just like the period of a pendulum. At growth rate 3.2, the system has a period-2 attractor. At growth rate 3.5, the system has a period-4 attractor. Just beyond growth rate 3.5, it bifurcates again into *eight* paths as the system oscillates over eight population values. These consecutive bifurcations are *phase transitions* from one behavior, such as a fixed-point attractor, to a qualitatively different type of behavior, such as a period-2 limit cycle attractor, as we vary the parameter value. Beyond a growth rate of 3.57, however, the bifurcations ramp up until the system is capable of eventually landing on any population value. This is known as the *period-doubling* path to chaos. As we adjust the growth rate parameter upwards, the logistic map will oscillate between two, then four, then eight, then 16, then 32 (and on and on to infinity) population values.

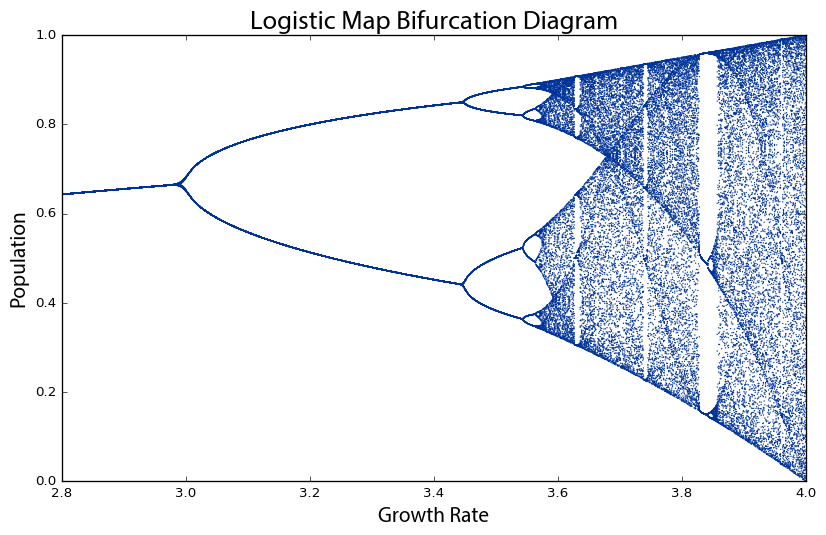


Figure 2.3. Bifurcation diagram of 100 generations of the logistic map for 1,000 growth rate parameter values between 2.8 and 4. The vertical slice above each growth rate depicts the system’s attractor at that rate.

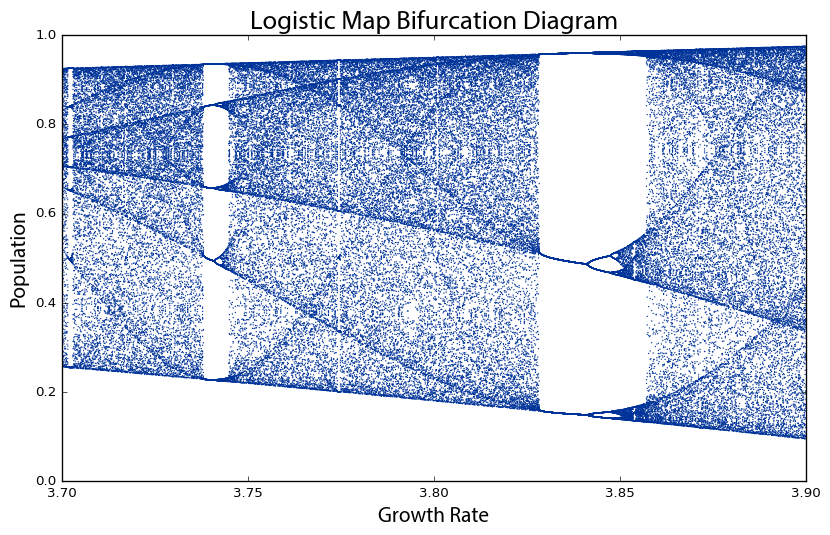


Figure 2.4. Bifurcation diagram of 100 generations of the logistic map for 1,000 growth rate parameter values between 3.7 and 3.9. The system moves from order to chaos and back again as the growth rate is adjusted.

By the time we reach growth rate 3.99, it has bifurcated so many times that the system now jumps, seemingly randomly, between all population values. We only say *seemingly* randomly because it is definitely *not* truly random. Rather, this model follows very simple deterministic rules yet produces apparent randomness due to its attractor having a period of infinite length. This is chaos: deterministic and aperiodic. If we zoom in again, to the narrow slice of growth rates between 3.7 and 3.9 (Figure 2.4), we begin to see the visceral beauty of chaos. Out of the noise emerge strange swirling patterns and thresholds on either side of which the system behaves very differently. For example, between the growth rates of 3.82 and 3.84, the system moves from chaos back into order, oscillating between just three population values: approximately 0.15, 0.55, and 0.95. But then at growth rates beyond 3.86 it bifurcates again and returns to chaos. Indeed *any* one-dimensional system with a period-3 cycle such as this at some parameter value is capable of chaotic behavior at other parameter values (T.-Y. Li & Yorke, 1975).

*Universality* refers to the phenomenon that very different systems can exhibit very similar behavior regardless of their underlying dynamics. It is commonly associated with Mitchell Feigenbaum’s discovery that all systems that undergo this period-doubling path to chaos obey a mathematical constant (Feigenbaum, 1978, 1983): the distance between consecutive bifurcations along the horizontal axis shrinks by a factor that asymptotically approaches 4.669, now known as *Feigenbaum’s constant* (Strogatz, 2014). Regardless of the system’s specific dynamics, the ratio of the bifurcations on its road to chaos always obeys this constant.

## Fractals and Strange Attractors

There is also a deep and universal connection between chaos and fractals (Tomida, 2008). In Figure 2.4, the bifurcations around growth rate 3.85 may look familiar. If we zoom in to the center one (Figure 2.5), we incredibly see the same structure that we saw earlier at the macro-level. In fact, if we keep zooming infinitely in to this visualization, we will continue seeing the same structure and patterns at finer and finer scales, forever. How can this possibly be? We mentioned earlier that chaotic systems have *strange attractors* and their structure can be characterized as *fractal* (Farmer, Ott, & Yorke, 1983; Grassberger & Procaccia, 1983; Hénon, 1976). Fractals are shapes that are self-similar, meaning they have the same structure at every scale (Mandelbrot, 1967, 1983, 1999). As we zoom in on them, we find smaller copies of the larger macro-structure. The bifurcation diagram (and thus the attractor) of the logistic map is a fractal: at the fine scale in Figure 2.5, we see a tiny reiteration of the same bifurcations, chaos, and limit cycles we saw in Figure 2.1’s visualization of the full range of growth rates.

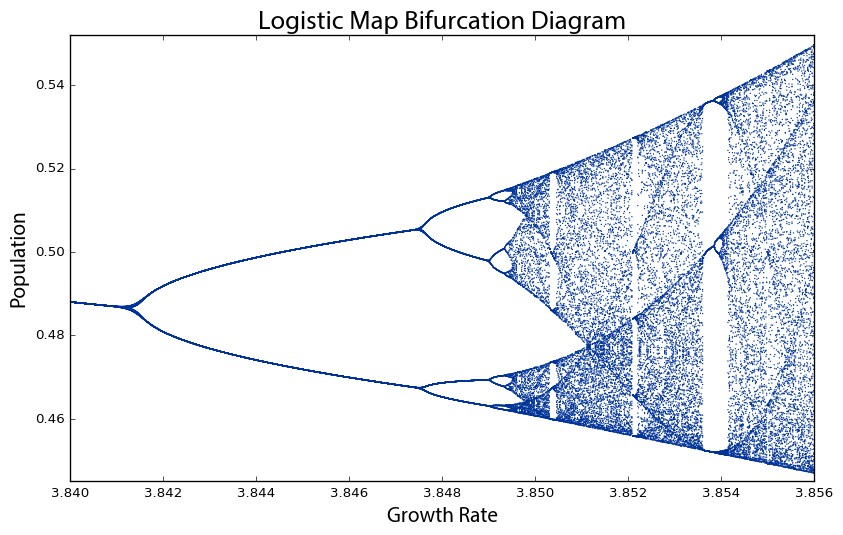


Figure 2.5. Bifurcation diagram of 100 generations of the logistic map for 1,000 growth rate parameter values between 3.84 and 3.856. This is the same structure that we saw earlier at the macro-level in Figure 2.3, because chaotic systems’ strange attractors are fractal.

Another way to visualize this nonlinear time series is with a *phase diagram*, using a method called state-space reconstruction through delay-coordinate embedding (Bradley & Kantz, 2015). Simply put, this plots the system’s value at generation *t*+1 on the *y*-axis versus its value at *t* on the *x*-axis (Huikuri et al., 2000), giving us another visual window into the qualitative behavior of the system. The clever insight of this phase diagram is that it embeds one-dimensional time series data from our logistic map into two-dimensional *state space*: an imaginary space that uses system variables as its dimensions (Packard et al., 1980; Takens, 1981; Theiler, 1990). Each point in state space is a system *state*, or in other words, a set of variable values. While traditional systems analysis tends to focus on visualizing time series as in Figure 2.1, nonlinear dynamics tends to focus on visualizing these state spaces. Few real-world systems are fully observable, yet the dynamics in a properly reconstructed state space are identical to the true dynamics of the entire system (Bradley, 2003).

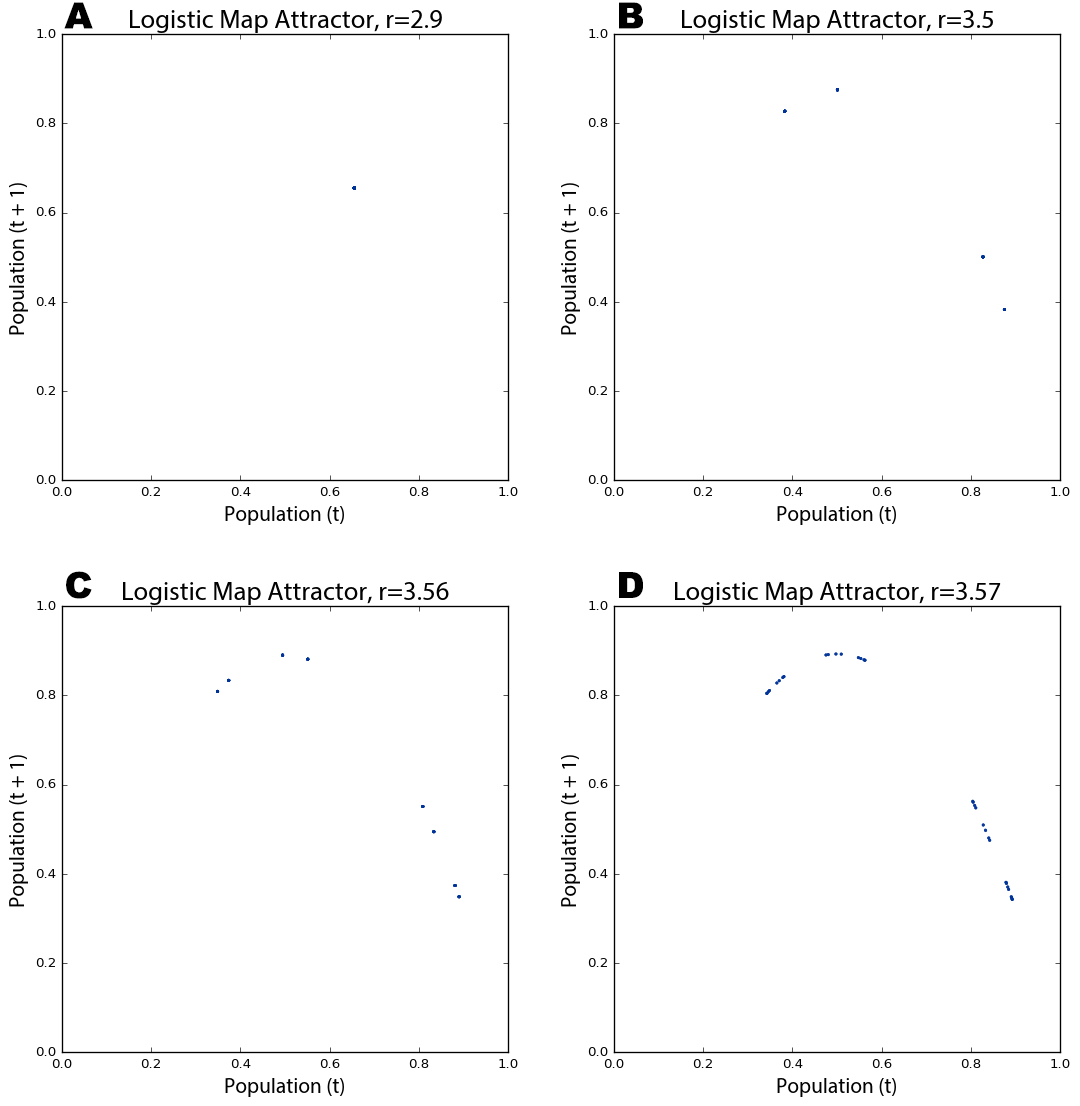


Figure 2.6. Phase diagrams of the logistic map over 200 generations for growth rate parameter values of 2.9 (A), 3.5 (B), 3.56 (C), and 3.57 (D). When the parameter is set to 2.9, the model converges at a single fixed-point. When the parameter is set to 3.5 or higher, the model oscillates over four points, then eight, and on and on as it bifurcates.

In our case, the two variables are 1) the population value at generation *t*, and 2) the value at *t*+1. For example, with a growth rate of 3.5, the population value at generation 1 is 0.5, the value at generation 2 is 0.875, the value at generation 3 is 0.383, and so forth (see Table 1). Therefore our two-dimensional phase diagram will have (*x*, *y*) points at (0.5, 0.875) and (0.875, 0.383) and so on (Figure 2.6b). Remember that our model follows a simple deterministic rule, so if we know a certain generation’s population value, we can easily determine the next generation’s value. Like earlier, to produce these phase diagrams Pynamical runs the logistic model for 200 generations and then discards the first 100 rows, to visualize only those values that the system settles toward over time.

In Figure 2.6a, the phase diagram shows that the logistic map homes in on a fixed-point attractor at 0.655 (on both axes) when the growth rate parameter is set to 2.9. This corresponds to the vertical slice above the *x*-axis value of 2.9 in the bifurcation diagram in Figure 2.2. Figure 2.6b depicts a period-4 limit cycle attractor: when the growth rate is set to 3.5, the logistic map oscillates over four points, as shown in this phase diagram (and in Figures 2.1 and 2.2). If we adjust the growth rate parameter up to 3.56, we witness a period-doubling bifurcation: Figure 2.6c shows the system now oscillating over eight points. As we approach the *chaotic regime* – the range of parameter values in which our system behaves chaotically – the period-doubling bifurcations start to come more quickly. Figure 2.6d shows that several additional bifurcations occurred between the growth rates of 3.56 and 3.57.

A kind of structure is slowly being revealed across Figure 2.6, but we can see it much more clearly as we push the growth rate parameter value deep into the chaotic regime. The phase diagram in Figure 2.7a reveals the system’s attractor at a growth rate of 3.9. Figure 2.7b visualizes 50 different growth rate parameter values between 3.6 and 4, each with its own color. Those rates that exhibit chaos form parabolas, but some gaps exist where the system occasionally settles down into periodic behavior (e.g., in the teal band when the growth rate is set to 3.83 – compare this band of periodicity with Figure 2.4).

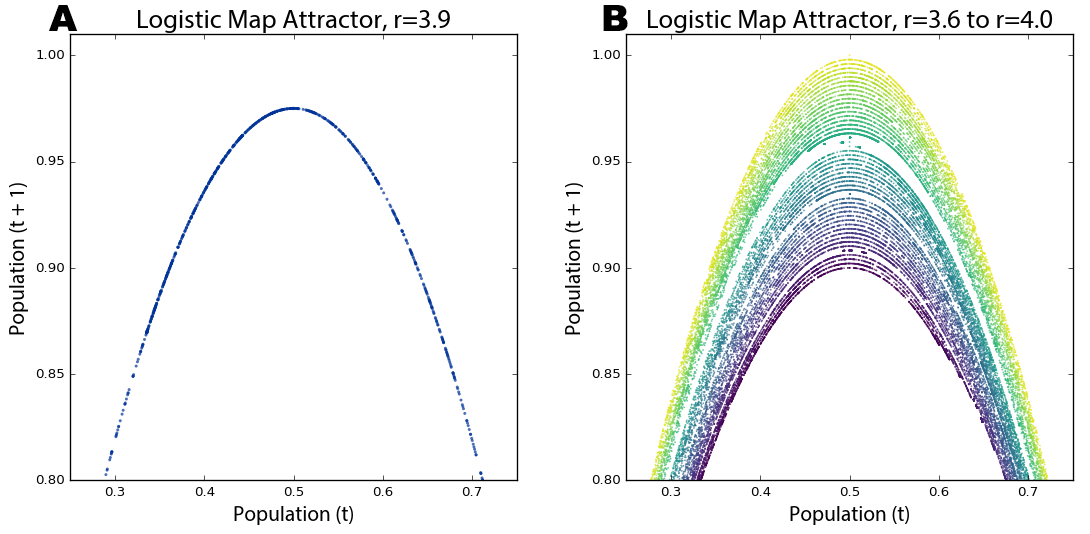


Figure 2.7. Cropped phase diagrams of the logistic map over 200 generations for (A) a growth rate parameter value of 3.9 and (B) 50 growth rate parameter values between 3.6 and 4 (the chaotic regime), each with its own colored line.

Strange attractors are revealed by these shapes as the system is somehow oddly constrained, yet never settles into a fixed point or limit cycle like it did in Figure 2.6. Instead it just bounces around different population values (i.e., points on the parabola) forever without ever repeating the same value twice. It is impossible to predict if any two consecutive observations appear near each other or far apart on the parabola. Further, the parabolas in Figure 2.7b never overlap due to their fractal geometry and the deterministic nature of the logistic map. Consider: if two different parameter values *could* ever land on the exact same point, their systems would have to evolve identically over time because the logistic map is deterministic. We can see in these visualizations that this indeed never happens. While the dynamics of a chaotic system appear to have no pattern whatsoever, in reality they conform to a remarkable fractal pattern – a strange attractor – which confines the system to a limited slice of state space and ensures that no state will ever repeat (Kekre, Sarode, & Halarnkar, 2014). Fractals are indeed strange. Rather than having a whole-number dimension such as two or three, they are characterized by a fractional (hence, fractal) dimension (Clarke, 1986; Grassberger & Procaccia, 1983; Theiler, 1990). The *fractal dimension* refers to the space-filling characteristics of a curve that, through self-similarity, becomes a bit more than a one-dimensional line yet a bit less than a two-dimensional plane.

These visualizations have all plotted quantitative data to better explain and understand the qualitative behavior of a nonlinear dynamical system. A *cobweb plot* is a visualization technique particularly well-suited to revealing the qualitative behavior of one-dimensional maps, allowing us to analyze the long-term evolution of such systems under recursive iteration (Tomida, 2008). The cobweb plots drawn by Pynamical in Figure 2.8 consist of three lines: a diagonal gray identity line representing *y*=*x*, a red curve representing the logistic map as *y*=f(*x*) for a given parameter value, and a blue line tracing the path of the cobweb. To draw a cobweb:

1. Begin on the *x*-axis at the point (*x*, 0) where *x* is the initial population value (0.5 in our example) and draw a vertical line to the red function curve – this new point is at (*x*, f(*x*)).
2. Draw a horizontal line from this point to the gray identity line – this new point is at (f(*x*), f(*x*)).
3. Draw a vertical line from this point to the red function curve – this new point is at (f(*x*), f(f(*x*))).
4. Repeat steps 2 and 3 recursively. The cobwebs in Figure 2.8 iterated 100 times.

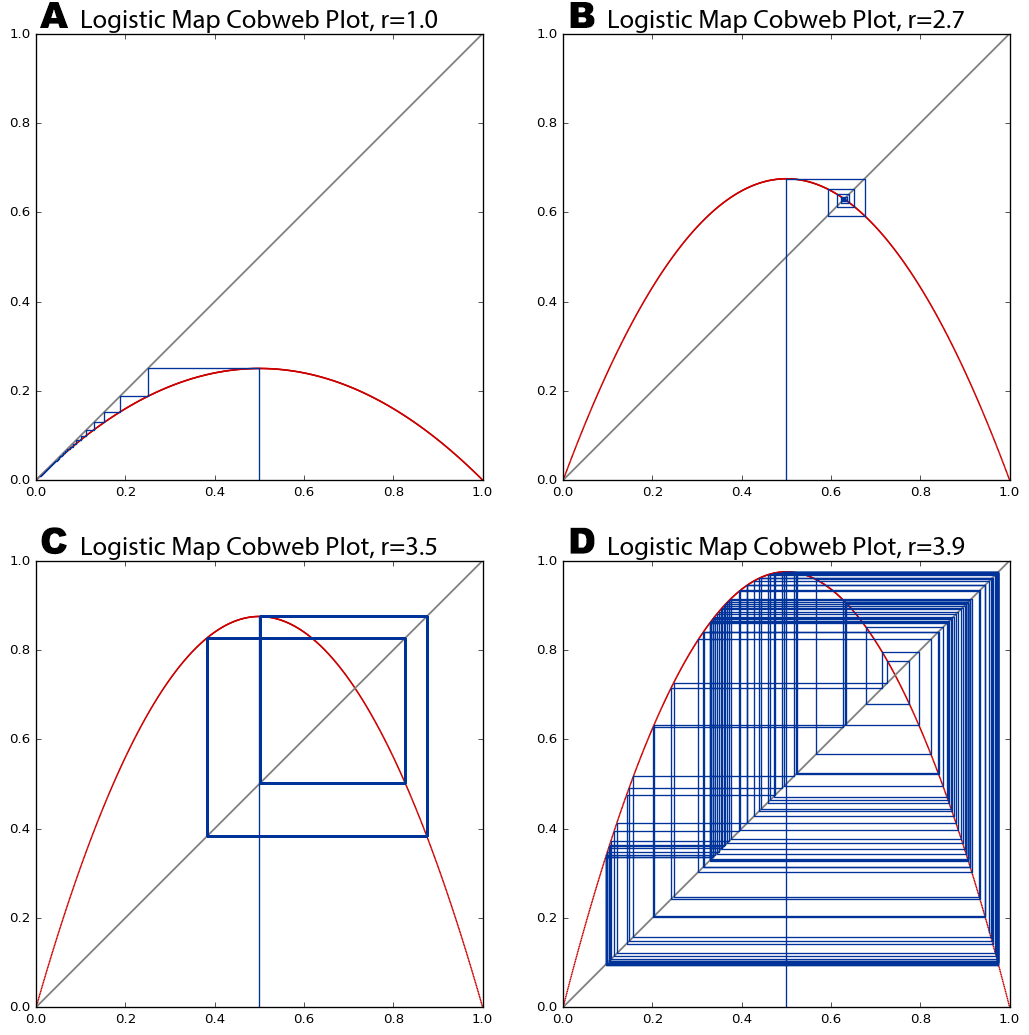


Figure 2.8. Cobweb plots of the logistic map for growth rate parameter values of (A) 1, (B) 2.7, (C) 3.5, (D) 3.9. The diagonal gray identity line represents *y*=*x*, the red curve represents the logistic map as y=f(*x*) for each of the four parameter values, and the blue cobweb line represents the system’s trajectory over 100 generations.

The blue lines intersect the red curve at those values our system lands on as it iterates from an initial population value of 0.5. In Figure 2.8a and b, the cobweb shows the system homing in on fixed-point attractors of 0 and 0.65, respectively. At a growth rate of 3.5 (Figure 2.8c) the system oscillates over four points in its limit cycle attractor, denoted by rectangular closed loops. The points where the blue lines intersect the red curve are the same as those revealed by the attractor in Figure 2.6b for the same parameter value. Finally, Figure 2.8d visualizes our system’s behavior in the chaotic regime at a growth rate of 3.9. The chaotic orbit fills the plot with rectangles – an eventually infinite number of never-repeating trajectories that form a fractal cobweb throughout the diagram.

## Chaos and Randomness

Phase diagrams are useful for visually revealing strange attractors in time series data, like that produced by the logistic map, because they embed this one-dimensional data into a two- or even three-dimensional state space. It can be difficult to ascertain if certain time series are deterministic or just random if we do not fully understand their underlying dynamics (Sander & Yorke, 2015). Take the two series plotted by Pynamical in Figure 2.9 as an example. Both of the lines seem to jump around randomly. The red line *does* depict random data, but the blue line comes from our logistic model when the growth rate is set to 3.99. This is deterministic chaos, but it is difficult to differentiate from randomness. Instead in Figure 2.10 we visualize these same two data sets with phase diagrams rather than time graphs, giving us a clear window into the qualitative behavior of our systems. Now we can clearly see our chaotic system constrained by its strange attractor. By contrast, the random data just look like the noise that it actually is.

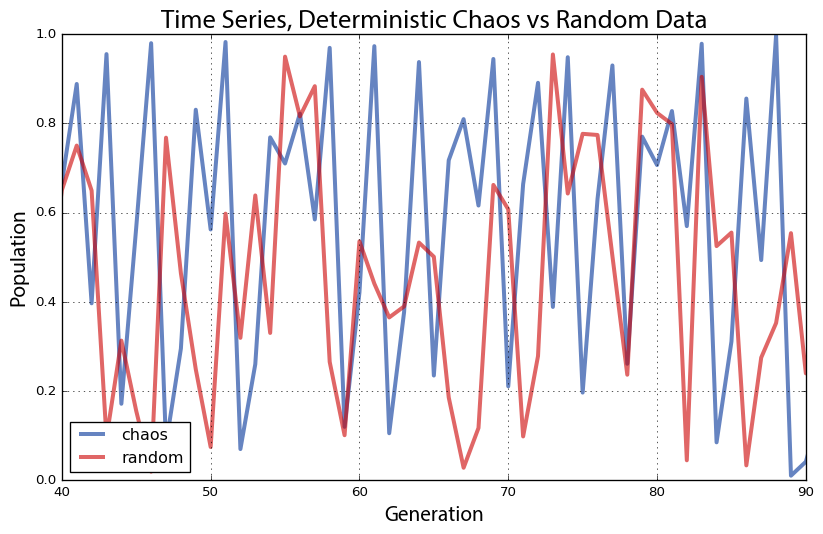


Figure 2.9. Plot of two time series, one deterministic/chaotic from the logistic map (blue), and one random (red).

This is particularly revealing in a three-dimensional phase diagram from Pynamical (Figure 2.10b) that embeds our time series into a three-dimensional state space by plotting the population value at generation *t*+2versus the value at *t*+1 versus the value at *t*. This plot essentially extrudes our two-dimensional plot (Figure 2.10a), then pans and rotates the viewpoint. In fact, if we looked straight down at the x-y plane of the three-dimensional plot in Figure 2.10b, it would look identical to the two-dimensional plot in Figure 2.10a (see Appendix for an animated visualization of this). Strange attractors stretch and fold state space in higher dimensions, allowing their fractal forms to fill space without ever producing the same value twice.

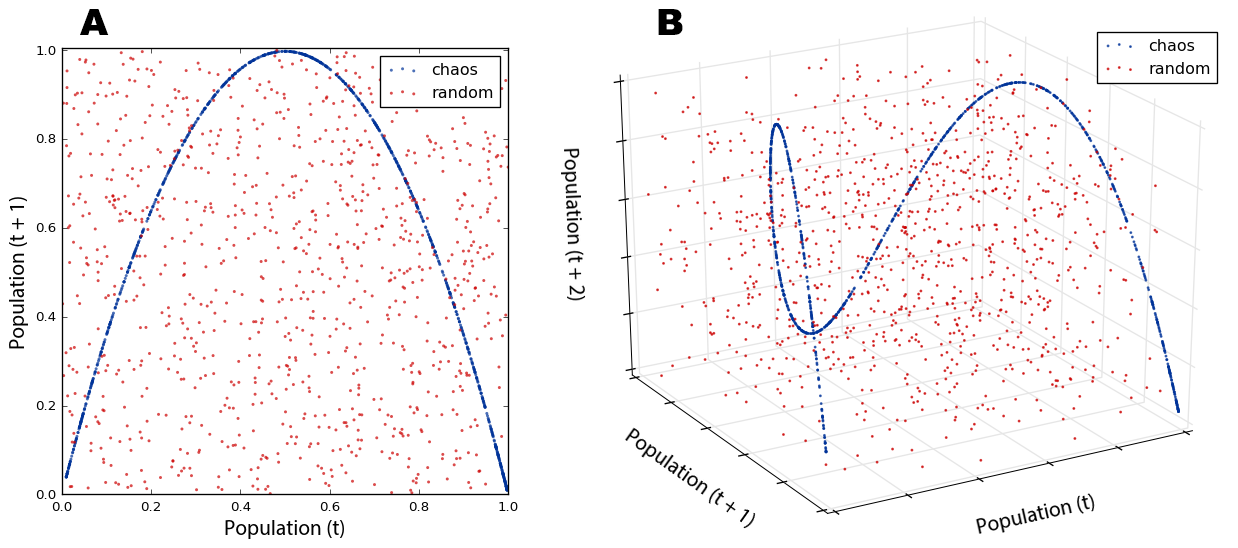


Figure 2.10. Phase diagrams of the time series in Figure 2.9. 10B is a three-dimensional state space version of the two-dimensional 10A.



Figure 2.11. Two different viewing perspectives of a single three-dimensional phase diagram of the logistic map over 200 generations for 50 growth rate parameter values between 3.6 and 4, each with its own colored line.

To press this further, we can use Pynamical to visualize the rest of the logistic map’s chaotic regime in three dimensions: the phase diagram in Figure 2.11 is a three-dimensional version of the two-dimensional state space we saw in Figure 2.7b. The color coding exposes the dynamical system’s behavior across the chaotic regime – information virtually impenetrable without visualization. The beautiful structure of the strange attractor is revealed as it twists and curls around its three-dimensional state space (see Appendix for an animated visualization). This structure again demonstrates that our *apparently* random time series data from the logistic model is not truly random at all. Instead, it is aperiodic deterministic chaos, constrained by a mind-bending strange attractor. No matter how much we zoom in, the parabolas never overlap and no point ever repeats itself.

## The Butterfly Effect

Attractors have a *basin of attraction*: a set of points that the system’s dynamics will pull into this attractor over time (Sprott & Xiong, 2015). This is easily seen with a cobweb plot. Figure 2.12 shows how the logistic map’s basin of attraction (when the growth rate is 2.7) pulls three different initial population values into the same fixed-point attractor. The initial state of the system will eventually become unknowable, because any one of many different possible points in the basin of attraction could have been the one pulled into the attractor.

By contrast, chaotic systems are characterized by their sensitive dependence on initial conditions. Their strange attractors are globally stable yet locally unstable: they have basins of attraction, yet within a strange attractor infinitesimally close points diverge over time without ever leaving the attractor’s confines. This divergence can be measured by *Lyapunov* *exponents* (Brown, 1996), the calculation of which is described by Wolf et al (Wolf, Swift, Swinney, & Vastano, 1985). If the Lyapunov’s value is positive, then the two points move apart over time at an exponential rate. If the Lyapunov is negative, then these points converge exponentially quickly, such as toward a fixed point or limit cycle. Finally, the Lyapunov is zero when there is a bifurcation (Dingwell, 2006). For example, with our logistic model, the Lyapunov is zero when the growth rate is set to 1 or 3 because they are bifurcation points; it is negative for most growth rates, such as 0 ≤ *r* < 1 and 1 < *r* < 3, because they have fixed-point or limit cycle attractors; and it is positive for the chaotic regime (exclusive of those occasional windows when the system resumes brief periodicity, such as when the growth rate is 3.83). A positive Lyapunov indicates that the system has a highly sensitive dependence on initial conditions, and is a common signature of chaos (Chan & Tong, 2013; Hunt & Ott, 2015; Kantz, Radons, & Yang, 2013).

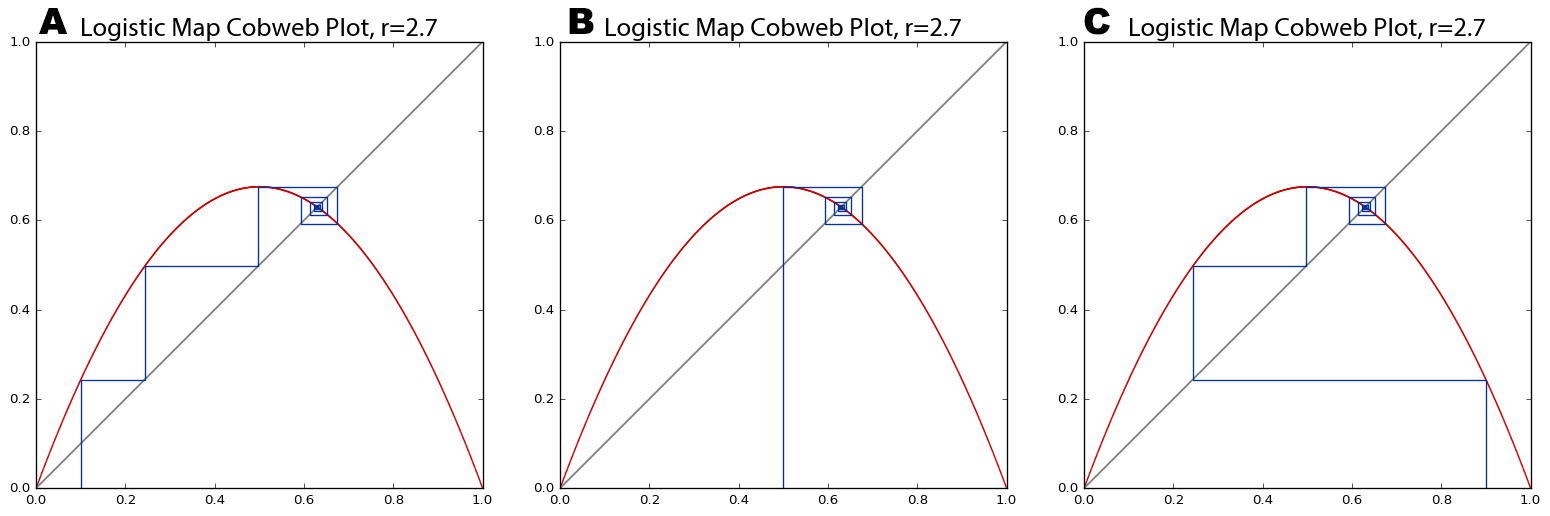


Figure 2.12. Cobweb plots of the logistic map pulling initial population values of 0.1 (A), 0.5 (B), and 0.9 (C) into the same fixed-point attractor over time. At this growth rate parameter value of 2.7, the Lyapunov is negative.

This nonlinear divergence of *very* similar values makes real-world modeling and prediction difficult, because we must measure the parameters and system state with infinite precision. Otherwise, tiny errors in measurement or rounding are compounded over time until the system eventually diverges drastically from the prediction. In the real world, infinite precision is impossible. It was through one such rounding error that Lorenz first discovered chaos. Recall his words at the beginning of this chapter: “the present determines the future, but the approximate present does not approximately determine the future” (Danforth, 2013).

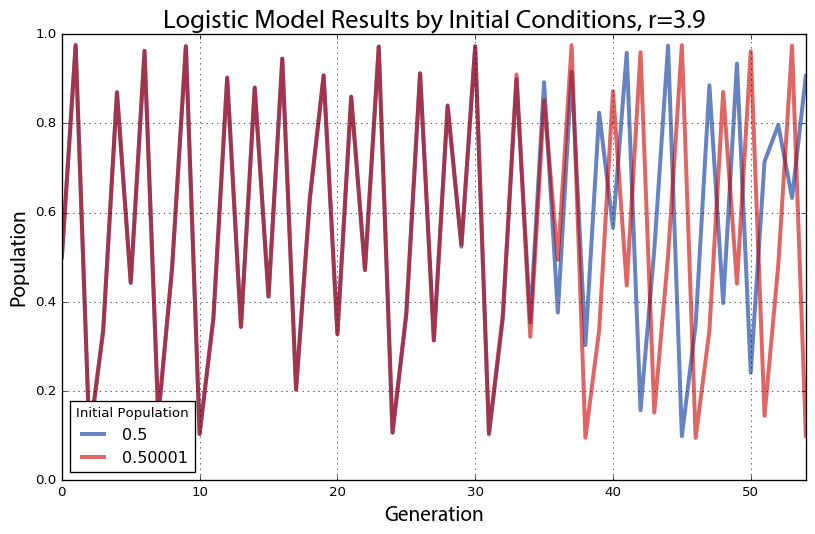


Figure 2.13. Plot of two time series with identical dynamics, one starting at an initial population value of 0.5 (blue) and the other starting at 0.50001 (red). At this growth rate parameter value of 3.9, the Lyapunov is positive – thus the system is chaotic and we can see the nearby points diverge over time.

As a demonstration of this, we run the logistic model with two *very* similar initial population values, shown in Figure 2.13. Both have the same growth rate parameter value of 3.9. The blue line represents an initial population value of 0.5 and the red line represents an initial population of 0.50001. These two initial conditions are extremely close to one another and accordingly their trajectories look essentially identical for the first 30 generations. After that, however, the minuscule difference in initial conditions compounds to the point that by the 40th generation the two lines show little in common. What began as nearly indistinguishable initial conditions produces completely different outcomes over time due to nonlinearity and exponential divergence.

If our knowledge of these two systems began at generation 50, we would have no way of guessing that they were nearly identical in the beginning. With chaos, history is thus lost to time and prediction of the future is only as accurate as our measurements. Human measurements are never infinitely precise, so in real-world chaotic systems, errors compound and the future becomes entirely unknowable given long enough time horizons. This phenomenon is famously known as the *butterfly effect*: a butterfly flaps its wings in China and sets off a tornado in Texas. Small events compound and irreversibly alter the future of the universe. In Figure 2.13, a tiny fluctuation of 0.00001 makes an enormous difference in the behavior and state of the system 40 generations later. Although this system’s future cannot be predicted, we *can* characterize its dynamics geometrically with phase diagrams, bifurcation plots, and cobweb plots – and statistically with Lyapunov exponents and fractal dimensions.

## Discussion

This chapter had two primary aims. First, it introduced the foundational concepts of nonlinear dynamics, chaos, fractals, self-similarity, and the limits of prediction through several visualization methods to analyze and understand the behavior of nonlinear dynamical systems. Second, it presented Pynamical, a software package for visualizing the behavior of discrete nonlinear dynamical systems. This package provides a free, fast, simple, extensible tool to introduce and analyze nonlinear dynamical systems’ behavior visually – useful for research and pedagogy. Nonlinear systems are extremely difficult to solve analytically because they cannot be broken down into constituent parts. Instead, we used Pynamical to reveal hidden structure and patterns in time series whose underlying dynamics may not be well known. In particular, it revealed the qualitative behavior of nonlinear dynamical systems over time and in response to parameter variations.

This chapter used the logistic map to define such a set of nonlinear dynamics. As simple as this model was, at different growth rate parameter values it produced stability, periodic oscillations, or chaos. We used Pynamical to create bifurcation diagrams and cobweb plots to visualize this behavior across different parameter values. In the chaotic regime, the system jumped seemingly randomly between all population values. Accordingly, we used Pynamical to embed the data into higher-dimensional state space to create phase diagrams to visualize the system’s strange attractor and understand its constrained, deterministic dynamics. Finally, we explored the butterfly effect’s implications of nonlinearity on system sensitivity, as infinitesimal differences in initial conditions compounded over time until nearly identical systems had diverged drastically. Thus in many nonlinear systems, there are fundamental limits to knowledge and prediction.

Pynamical and all of the code used to develop these models and produce these visualizations are available in a public repository on GitHub at https://github.com/gboeing/pynamical. Pynamical is built on top of Python’s pandas, numpy, and matplotlib code libraries:

* numpy is a numerical library that handles the underlying numerical vectors
* pandas handles the higher-level data structures and analysis
* matplotlib is the engine used to produce the visualizations and graphics

Pynamical defines extensible functions to express the discrete map’s equation and encapsulate the model that runs the equation iteratively. The logistic map, the Singer map, and the cubic map are built-in by default but any other iterated map can be defined and added. Pynamical also defines a function to convert model output into x-y points, as well as functions to plot these points as a bifurcation diagram, a cobweb plot, an animated cobweb plot, a two-dimensional phase diagram, a three-dimensional phase diagram, and an animated three-dimensional phase diagram. Animated cobweb plots of the entire parameter space and animated three-dimensional phase diagrams extending those presented in this study are also available in the GitHub repository. They shed particular light on the fractal nature of strange attractors as they stretch and fold state space, thus serving as an indispensable tool for pedagogy and visual information presentation.

Pynamical is very simple to use and serves as a useful tool for introducing nonlinear dynamics and chaos. Sample code to produce some of the visualizations in the chapter demonstrates this simplicity. One merely imports Pynamical into the Python environment then runs the following code to produce the visualization:

* Figure 2.2: bifurcation\_plot(simulate(num\_rates=1000))
* Figure 2.4: bifurcation\_plot(simulate(min=3.7, max=3.9, num\_rates=1000))
* Figure 2.6d: phase\_diagram(simulate(num\_gens=100, min=3.57))
* Figure 2.8d: cobweb\_plot(r=3.9, x0=0.5)
* Figure 2.11: phase\_diagram\_3d(simulate(num\_gens=4000, min=3.6, num\_rates=50)

Python was selected for developing Pynamical and the visualizations in this chapter because it is a standard programming environment for information visualization. Python is fast, free, powerful, and open-source. Although it is an interpreted language, several of its libraries such as numpy provide compiled components for vectorized numerical functions, making mathematical modeling fast and efficient. These can take advantage of the optimized mathematical routines of Intel processors’ Math Kernel Library. Python is multi-purpose and a researcher can use its standard syntax and grammar for everything from statistical modeling, to cartography, to full software development. Becoming a Swiss army knife of the computational science world, Python has grown popular and powerful. Today innumerable researchers and developers contribute open-source libraries of pre-packaged functionality for all to repurpose.

## Conclusion

During the 1990s, complexity theories grew out of chaos theory and largely supplanted it as an analytical frame for social systems. Although complexity draws on similar principles, it emerges as a very different beast. Instead of looking at simple, closed, deterministic systems, complexity examines large open systems made of many interacting parts. Unlike chaotic systems, complex systems retain some trace of their initial conditions and previous states, through path dependence. They are unpredictable, but in a different way than chaos is: complex systems have the ability to surprise through novelty and emergence. The following chapter builds on this foundation of nonlinearity to unpack the theory of complex systems, examine how it applies to the study of cities, and introduce complex spatial network analysis.

# Complexity and Cities

## Abstract

This chapter presents the theoretical framework of complex systems and cities, culminating in network theory and analysis – the primary lens this study uses in all subsequent chapters. Discussions of complexity and complex systems have appeared throughout the planning literature for years. These principles have been applied everywhere from the communicative turn and collaborative rationality, to cellular automata and agent-based urban models, to the design of resilient, livable neighborhoods. However, the interdisciplinary appeal and trendiness of complexity in the social sciences has resulted in a bit of a morass of ambiguous terminology, internal. This chapter unpacks the key foundational concepts of complex systems and network science in a brief, straightforward manner. It provides explanatory examples of these concepts familiar to scholars and practitioners not already versed in the technical science of complexity. Finally, it outlines the key implications this interdisciplinary science offers to the scholarly field of urban planning and the real-world practice of city-making.

## Introduction

Complexity theories have become a popular frame for conceptualizing and analyzing cities. There is no single complexity *theory* but rather a wide array of concepts and tools that can be applied to the study of complex systems across numerous disciplines (Manson and O’Sullivan 2006; Haken 2012). The term *complexity theory* generally refers to a nebulous union of these theories. Such theories propose that certain large systems are characterized by the nonlinear, dynamic interactions of their many constituent parts. These systems then behave in novel and unpredictable ways—ways that cannot be divined by simply examining the components of the system. Complexity problematizes traditional reductionist, linear methods of scientifically analyzing and predicting cities. It also opens up a new world of scholarship to researchers keen to formulate new kinds of sciences that take complexity into account (e.g., Wolfram 2002). These attempts usually follow Kuhn’s (1962) theory of paradigm shifts: new evidence and modes of thinking undermine an established science, and a new science emerges to replace it.

Complexity theories have become a popular framework for scholarly enquiries into planning and urban studies over the past 30 years. Although it entails a fundamental shift away from the belief that predictive certainty is possible with complex systems, it *can* serve as a useful new lens for explaining urban phenomena, studying city form, and considering planning interventions. Further, complexity provides a comprehensive framework that might build connections between quantitative and qualitative urban disciplines (Portugali 2006). However, complexity theories are sometimes adopted into the social science literature in obscure or contradictory ways.

In particular, complexity theory in the planning literature has suffered from three notable problems. First, physical scientists often apply it atheoretically, either unconstructively problematizing planning methods, or naively (but mathematically) “proving” long-recognized urban phenomena. Second, planning scholars sometimes cherry-pick concepts from complexity then use them abstractly or vaguely. For example, Portugali (2012) has criticized planning theorists like Manuel Castells and Patsy Healey for borrowing from complexity theory, then using the idea of complexity merely vernacularly – thereby losing all of its formalism and implications. Chettiparamb (2006) has critiqued David Byrne for relying on undefined jargon, obscure phrases, and contextless assertions that make otherwise-scientific complexity theory feel vague and mystical. The interdisciplinary appeal and trendiness of complexity in the social sciences has resulted in a bit of a morass of ambiguous terminology, internal inconsistencies, and overloaded concepts open to multiple interpretations. Third, because of the first two problems, some urban scholars have dismissed “complexity” as merely fashionable nonsense.

This chapter wades into this morass to unpack the essential shared, foundational concepts of complexity theories – particularly as they might apply to cities – in a brief, straightforward manner. It is organized as follows. First, it draws the nonlinear foundations from chapter 2 out of the realm of simple closed systems and into the world of complex open systems. Then it discusses measures of complexity, providing a framework that will be fleshed out in applied detail in chapter 4. Next it reviews the concepts and ramifications of equilibrium, stability, emergence, self-organization, and resilience, drawing these concepts from the natural sciences into the study of cities, particularly urban form. This leads to the chapter’s final section, in which it presents the science of networks and network analysis, laying the theoretical foundation for chapters 5 and 6.

## Systems and Dynamics

A system is a set of interacting components that together form a whole. In the context of complexity theory, to say that a system is complex is to say that we cannot understand its behavior simply by examining its constituent parts (Mitchell 2009). Complex systems comprise many interacting subcomponents whose recurrent interactions cause nonlinear feedback, collective behavior, and unpredictable emergent phenomena at larger scales (Rickles et al. 2007). In contrast, complicated refers merely to being made up of many interrelated parts.

Examples are useful to disambiguate these types of systems. A wind-up clock is an example of a simple system with few interrelated parts. An automobile is an example of a complicated system with many interrelated parts. In contrast, stock markets, the climate, and cities are examples of complex systems. Complex systems are defined more by their internal relationships than they are by their constituent parts (Manson 2001). It is this networked structure and organization that makes them interesting. The term complexity itself refers to the rich, dynamic system behavior arising from individual interactions between many system subcomponents.

As discussed in chapter 2, a dynamical system changes over time as its state evolves according to its initial conditions and the processes that describe its subcomponents’ behavior. A system’s state is the essential information about the system for an observer and is defined by the values of the relevant variables. A variable could be a system feature or a calculated indicator that an observer has decided to use to describe the system. Process is more difficult to define, but generally refers to some sequence of actions that changes the system state (O’Sullivan and Perry 2013). Related to process, dynamics can be interchangeably thought of as the system’s rules or, thus, the paths the system state traces through time. These paths of the variables through time are visualized with phase space diagrams, as discussed in chapter 2. Phase space is an abstract space that contains all possible system states, with each possible state represented by a single point.

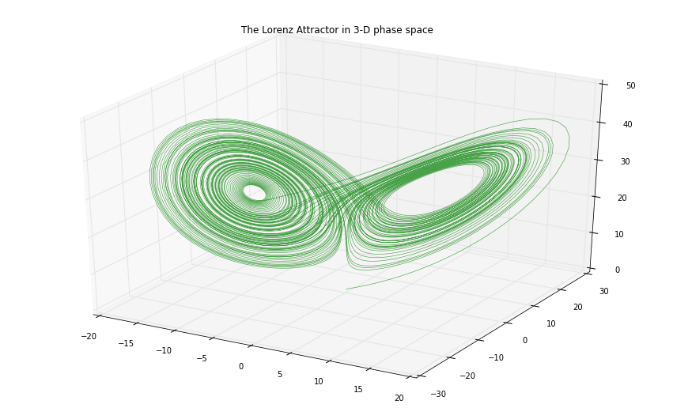


Figure 3.1 An example phase space diagram of the evolution of the Lorenz system over time.

Real-world complex systems are sensitive to outside influences because they are open systems. An open system is one that cannot be screened off from its environment, so researchers cannot safely ignore outside influences – or in other words, exogenous variables. Most real-world systems are open and pose problems for modeling because these exogenous influences must be taken into account (Batty and Marshall 2012). A model is simply an abstract representation of something, and could refer to conceptual models, mathematical models, statistical models, or even physical models (O’Sullivan and Perry 2013). Although real world systems tend to be open, models tend be closed to be tractable.

Complex systems and chaotic systems are both subtypes of nonlinear dynamical systems. Complexity theory deals with complex open systems that self-organize into emergent forms that could not have been predicted simply by understanding the constituent parts (Mitchell 2009). Chaos theory deals with simple, deterministic, nonlinear, dynamic, closed systems resulting in a chaotic response to different initial conditions or perturbations (Reitsma 2002). As discussed in chapter 2, chaotic systems are unpredictable beyond limited time horizons because of their sensitivity to initial conditions, and these initial conditions may become unknowable later. Complex systems, in contrast, are unpredictable because of their capacity for novelty via emergence, discussed in detail below. They are sensitive to initial conditions in the sense that early historical accidents create path dependence that maintains their legacy over long time horizons. Chaos theory examines apparent disorder arising from simple order, while complexity theory examines large-scale order emerging from disorder at the local scale (ibid.).

## Measures of Complexity

There are several measures of complexity. Mitchell (2009) points out that no one measure is ideal (or could possibly capture the myriad denotations and connotations of complexity), but highlights several prominent ones: information entropy from information theory; statistical complexity, or the degree of structure and pattern in a system; the Lyapunov exponent which mathematically defines a system’s sensitivity to initial conditions; and the fractal dimension, which defines complexity as the irregularity of an object’s form.

The former two are discussed in detail with regards to urban form and street networks in chapter 4. The latter two were explored in chapter 2 for simple nonlinear systems, but their implications for complex systems are worth considering briefly here. Sensitivity to initial conditions makes prediction of a nonlinear system difficult, as the initial state must be described with perfect accuracy (Rickles et al. 2007). Unfortunately, measurement of the real world always requires some amount of rounding and entails some amount of uncertainty. These tiny inaccuracies compound over time as the system evolves, making prediction difficult or perhaps impossible. Theorists from Friedrich von Hayek (1974) to Ilya Prigogine (1997) have thus questioned whether it is even possible to make accurate predictions of complex systems, given the requirements of data-gathering and precision *and* because of such systems’ capability to surprise via emergence, a concept discussed in section 3.6.

Complex systems such as cities are sensitive to initial conditions in the sense of historical accidents, but their path dependence continues to reveal these conditions over long time horizons. Path dependence simply refers to the idea that history matters: complex systems’ past states are remembered and play a role in future states (i.e., they are non-Markovian systems). Further, it is possible for single events to alter a complex system in a way that persists for a long time. In cities, historical accidents/natural subsidies (i.e., sensitivity to initial conditions) or exogenous perturbations (e.g., wars, new technology, or economic shocks) may greatly affect long-term system behavior. Some echo of a complex system’s initial conditions remains apparent far into the future, whereas a simple chaotic system’s initial conditions are eventually lost to time and become unknowable.

Finally, as introduced in chapter 2, the fractal dimension refers to the non-integer dimension of an object with an irregular form – e.g., a line so kinked that it can be characterized as something between a one-dimensional line and a two-dimensional plane. Complex systems such as cities produce fractal self-similar forms that can be seen in urban peripheries and street networks (White and Engelen 1993; Batty and Longley 1994; Benguigui et al. 2000; Shen 2002) – and at the scale of urban design, as will be explored in chapter 4.

## Equilibrium and Stability

Equilibrium is used in different ways in the social sciences literature. In urban economics, equilibrium typically refers to a point at which supply and demand are balanced, resulting in – to use location choice as an example – no incentive for anyone to move (O’Sullivan 2008). However, complexity scholars typically borrow from physics instead to define equilibrium as a steady state of constant, maximum entropy in which a system does not change, adapt to its environment, or evolve structure. A common illustrative example is a gas diffusing into a vacuum until it is evenly dispersed. In the 1970s, Ilya Prigogine (Nicolis and Prigogine 1977) discovered that certain far-from-equilibrium open systems can evolve structures that locally contradict the second law of thermodynamics, which states that systems move toward maximum entropy. Allen and Sanglier (1981) extended Prigogine’s findings to the urban studies literature through their reformulation of central place theory in terms of these dissipative structures and bifurcation. Given these different definitions, there is some vagueness in how the term “equilibrium” is used, often unqualified, in the urban complexity literature.

Sometimes it refers to thermodynamic equilibrium, as scholars invoke it to argue that cities are far-from-equilibrium complex systems in the Prigogine sense. This stream of literature argues that cities are open systems and thus matter and energy – such as food, electricity, immigrants, building materials, etc. – flow into them. Structure and order evolve, locally violating the second law of thermodynamics. In this sense, cities do not move toward equilibrium; rather they are far from it, ever evolving and structuring their matter (White and Engelen 1993).

Other times, equilibrium is used to refer to an equilibrium of dynamics, where the system becomes limited and its state settles into an unchanging value or a set of values that it oscillates over. This is known as a stable equilibrium, or for disambiguation’s sake, a stable state. Equilibrium in this context means that a system is in balance despite multiple forces acting on it, dominated by negative feedback that damps perturbations and pushes the system back toward the equilibrium. At a macro-level, a real-world complex system might appear to be in static equilibrium based on its system-level state variables, but at a micro-level components may be dynamic and in flux. Consider residents settling into locations in a city. Over time, stable and consistent patterns may emerge city-wide, but at the human scale, residents are always moving in, around, and out of the city. Stable states make for tractable models of complex systems: Schelling (1971) famously demonstrated a simple simulation in which fairly tolerant residents relocate based on their subtle preferences for similar neighbors, resulting in a surprisingly segregated equilibrium.

Complex systems can settle for periods of time into stable, metastable, or unstable states – or even shift between alternative stable states via phase transitions. These are depicted by the balls in Figure 3. A stable state (ball 3) is one in which the system is resilient to perturbation and its dynamics return it to this state after being perturbed. Stable states may include steady states – in which the system state remains at some fixed value – or limit cycles – in which the system oscillates over a consistent set of values. An unstable state (ball 2) is one which the system moves *away* from after even a slight perturbation: the system is precariously perched between two possible states that it could settle into. A metastable state (ball 1) is one the system returns to after small but not large perturbations. The system may spend extended time in this state, but a sufficient perturbation will push it into a “preferred” state.

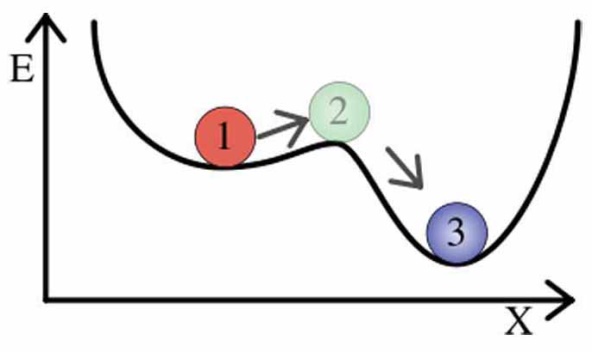


Fig. 3: Metastable (ball 1), unstable (ball 2), and stable (ball 3) states. Ball 3 will return to its current state after a large perturbation. Ball 2 will move away from its current state after even a very slight perturbation. Ball 1 will remain in its metastable state after a small perturbation, but a large one will push it into a preferred stable state. Source: http://www.psych.utah.edu/~jb4731/systems/Lexicon.html

Alternative stable states are also possible. Ecosystems can exist in different stable states over long periods of time. After a certain perturbation, they may transition from one to another via a phase transition (also known as a regime shift). Such behavior suggests a system with possible states that are separated by thresholds rather than a smooth gradient, a common outcome of nonlinearity. Hysteresis – the dependence of a system’s behavior on both its present state and its past states – allows a system to exist in different states at different times but under the same conditions. This path dependence helps keep the system in the current state and suppresses transitions to other states it could otherwise be in.

Criticality is related to metastability. But unlike shifting from one (fairly) stable state to another, criticality connotes a system poised on the edge of catastrophe. A system is critical if its behavior changes dramatically – for instance, transitioning from an ordered regime to a chaotic one – given some small input. This critical state was popularly referred to in the past as the “edge of chaos.”

When a parameter is adjusted to the critical point, the system undergoes a quick, radical change in its qualitative features. A parameter is a factor that defines the system. A model parameter is similar to a variable, but is either some universal value or is directly controlled by the researcher (rather than simply being observed). For the latter, consider the phase change of water at certain temperatures, from gas to liquid to solid. Here, temperature is the parameter being adjusted by the researcher to the critical point. Alternatively, consider parameters such as the CO2 carrying capacity of the Amazon rain forest, which may be some constant at any given time, but could change through global warming or pollution.

Finally, as discussed in chapter 2, bifurcation is the tendency for a system or one of its variables to jump suddenly from one attractor or stable state to another. When this happens, a drastically different nature of the system appears. Allen (2012) argues that bifurcation in complex human systems – such as cities – can be interpreted as a key historical moment in time when system could go one direction or another, with multiple possible future trajectories. With a sufficient understanding of the system and its dynamics, planners may adjust a parameter to shape the trajectory toward socially desirable outcomes.

## Emergence, Self-Organization, and Resilience

An emergent system property arises from interactions between subcomponents of a complex system. These subcomponents, however, do not themselves display this new system property and the property could not have been deduced merely by examining the subcomponents and their interactions (Aziz-Allaoui and Bertelle 2009). Nonlinearity is the source of emergent system characteristics, as they result from many repeated interactions among subcomponents that eventually exceed the sum of the parts (Manson 2001). In other words, the researcher cannot just take the system apart, inspect the components to understand what the system does, and them put them back together again. Emergent phenomena are nonlinear characteristics of a system, such as catastrophes, thresholds, and self-organization.

Self-organization is an emergent phenomenon that occurs when a system orders itself into a “better” or more stable state without external control or a central overseer. This tends to be a bottom-up process by which one hierarchical level generates the features of the level above it. The distinction between top-down and bottom-up processes should not be taken to be binary. Rather, each simply refers to the general directionality of a process in terms of the system’s hierarchy.

Feedback occurs when an output of the system “feeds back” into the system as an input. Negative feedback damps a variable’s rate of change and pushes it toward a stable state. Positive feedback increases a variable’s rate of change, as self-reinforcement. Furthermore, large-scale structures can emerge from small-scale subcomponent behavior and then influence future subcomponent behavior via cross-scale feedback (Allen 2012). Through co-evolution, subcomponents create their environment and are then in turn influenced by it. Culture, religion, and social norms – created by humans and in turn influencing humans – are examples of such emergent properties and their cross-scale feedback within urban agglomerations of humans.

Resilience and robustness are related to self-organization. Walker et al. (2004, p. 6) define resilience as “the capacity of a system to absorb disturbance and reorganize while undergoing change so as to still retain essentially the same function, structure, identity, and feedbacks – in other words, stay in the same basin of attraction.” Thus, a resilient system is able to return to its original stable state after a perturbation. Robustness, in contrast, tends to refer to the stability of a specific variable or system characteristics despite instability among some system components (Aligica 2014). To put it simply, resilience refers to returning to an original state after a perturbation, while robustness refers to perturbations having only a minimal effect in the first place.

Self-organized criticality describes a dynamical system that has a critical point as its attractor and continually evolves itself (without any external tuning or management) to this point of phase transition and catastrophe. At the critical point, system subcomponents are extremely connected and strongly influence each other. Here, even a small change to a single subcomponent is capable of producing vast effects that ripple through the entire system. The classic example is the sand pile model described by Bak et al. (1988). In it, a sand pile has additional grains of sand continuously dropped on top of it. The pile evolves to a certain angle of repose – the critical point – despite frequent trivially small avalanches. At this point, a single additional grain can suddenly cause a massive avalanche. After an avalanche, the system slowly evolves back to that critical angle and repeats. Batty and Xie (1999) argue that cities exhibit self-organized criticality as their urban forms evolve over time through discrete transitions. Forest fires are another example: frequent small fires tend to prevent fuel build-up, but huge fires are occasionally possible if no small fires have cleared the underbrush in a long time (Malamud et al. 1998).

Such systems accumulate energy over time and dissipate it through many small and few large events. Thus, systems that exhibit self-organized criticality produce events – such as sand pile avalanches, forest fires, and earthquakes – that range from tiny to enormous (Turcotte 1999). In other words, these systems are scale-free – they have no characteristic size. Human beings, on the other hand, have a characteristic size as they tend to range between five and seven feet tall, with few outliers. Self-similarity, scale invariance, scale-free, and fractal geometry are equivalent concepts that indicate a lack of this characteristic scale (Rickles et al. 2007). Scale free systems follow a scaling law such as a power law with the form *p=b-a*. Gaussian (normal) distributions result from processes that tend to sum to the center of the range, but with a power law distribution, the probability *p* of an event is inversely proportional to its size *b*. Thus, there are very few massive events, some medium-sized ones, and lots of small ones.

Similar characteristics can be seen in numerous systems despite their different underlying dynamics. This phenomenon is called universality. Accordingly, in the 1990s power laws became something of a popular signature for an underlying complex system. However, an over-reliance on power laws and universality has met considerable controversy and criticism in recent years (Stumpf and Porter 2012). It can be challenging to differentiate between a power law distribution and other candidate distributions, particularly lognormal if it is difficult to observe tiny events to the left of the mode peak. Further, there are innumerable ways to generate a power law distribution, so it alone cannot be an unambiguous indicator that a complex system underlies the observed phenomena (Mitzenmacher 2004).

This plays into a classic challenge of complex systems study: the equifinality problem is that different processes and models can result in the same outcome or pattern (Beven and Freer 2001). Many urban processes can be shown to produce similar patterns across spatial scales, but this does not help us understand what exactly is happening in each instance. For example, urban form may have a fractal spatial pattern, but this finding has yet to be connected convincingly to underlying social and economic processes (Manson and O’Sullivan 2006). Although much of the purpose of complexity studies lies in linking patterns to processes, there is great risk in conflating pattern *with* process.

## Networks

This chapter has thus far discussed complex systems theory, drawn from the foundations of nonlinear systems presented in chapter 2, including dynamics, stability, emergence, and self-organization. These characteristics appear in a system as a result of the many interactions between its connected, constituent parts. It is in fact these interactions and connections within a system that make up the backbone of its study. This is the subject of network science.

Network science is built upon the foundation of graph theory, a branch of discrete mathematics. A graph is an abstract representation of a set of elements and the connections between them (Trudeau 1994). The elements are interchangeably called vertices or nodes, and the connections between them are called links or edges. For consistency, I will use the terms *nodes* and *edges* throughout this study. The number of nodes in the graph (i.e., the *degree* of the graph) is commonly represented as *n* and the number of edges as *m*. Two nodes are *adjacent* if an edge connects them, two edges are adjacent if they share the same node, and a node and an edge are *incident* if the edge connects the node to another node. A node’s *degree* is the number of edges incident to the node, and its *neighbors* are all those nodes to which the node is connected by edges.

An *undirected* graph has undirected edges (i.e., each edge points mutually in both directions) but a *directed* graph, or digraph, has directed edges (i.e., edge *uv* points from node *u* to node *v*, but there is not necessarily an edge *vu*). A *self-loop* is an edge that connects a single node to itself. Graphs can also have parallel (i.e., multiple) edges between the same two nodes. Such graphs are called *multigraphs*, or *multidigraphs* if they are directed. An undirected graph is *connected* if each of its nodes can be reached from any other node. A directed graph is *weakly connected* if the undirected representation of the graph is connected, and *strongly connected* if each of its nodes can be reached from any other node. A *path* is an ordered sequence of edges that connects some ordered sequence of nodes. Two paths are *internally node-disjoint* if they have no nodes in common, besides end points. A *weighted* graph’s edges have a weight attribute to quantify some value, such as importance or impedance, between connected nodes. The *distance* between two nodes is the number of edges in the path between them, while the *weighted distance* is the sum of the weight attributes of the edges in the path.

While a graph is an abstract representation of elements and their connections, a *network* may be thought of as a real-world graph. Networks inherit the terminology of graph theory. Familiar networks include social networks (where the nodes are humans and the edges are their interpersonal relationships), the Internet (where the nodes are computers and the edges are the physical TCP/IP-based links that connect them), and the World Wide Web (where the nodes are web pages and the edges are hyperlinks that point from one to another). A complex network is one with a nontrivial *topology* (the configuration and structure of its nodes and edges) – that is, the topology is neither fully regular nor fully random. Most large real-world networks are complex (Newman 2010). Of particular interest to this study are *complex* *spatial networks* – that is, complex networks with nodes and/or edges embedded in space (O’Sullivan 2014). A street network is an example of a complex spatial network with both nodes and edges embedded in space, as are railways, power grids, and water and sewage networks (Barthelemy 2011).

A spatial network is *planar* if it can be represented in two dimensions with its edges intersecting only at nodes. A street network, for instance, *may* be planar (particularly at certain small scales), but most street networks are non-planar due to grade-separated expressways, overpasses, bridges, and tunnels. Despite this, most quantitative studies of urban street networks represent them as planar networks (e.g., Strano et al. 2013) for tractability because bridges and tunnels are (in some places) reasonably uncommon, and thus the networks are *approximately* planar. However, this over-simplification to planarity for analytical tractability may be unnecessary and can cause analytical problems, as we shall discuss in chapter 5.

Complex networks have been used extensively by urban scholars and planning researchers. From a qualitative perspective, Castells (e.g., 2009) argues that understanding flows and networks, rather than locations themselves, are the key to understanding and designing cities. From a quantitative perspective, Batty (e.g., 2013) places urban modeling in the context of network evolution and flow. Most relevant to this study, however, is rich body of transportation and urban form studies that use complex street networks for routing and characterizing the structure of cities. In particular, a typology of measures of the complexity of urban form – and particularly street networks – is developed in chapter 4 and applied empirically in chapters 5 and 6.

## Conclusion

Complex systems are systems of interacting components that through nonlinearity can create emergent phenomena and self-organized structure. Human society and cities are examples of large, complex systems. Principles of complexity have been applied in urban planning from the communicative turn and collaborative rationality, to cellular automata and agent-based urban models, to the design of livable neighborhoods. In particular, the emergent features of stability, resilience, robustness, and connectivity are of particular interest to urban scholars. The latter is a key bridge between qualitative theories of cities, such as Castells’ spaces of flows, and quantitative studies of the cities – broadly, the study of urban form, design, and transportation. The next chapter manifests these theories of complexity into the discipline of urban design and builds a typology of measures of the complexity outcomes of urban design, particularly emphasizing those relevant to street network analysis.

# Methods for Measuring the Complexity Outcomes of Urban Design

## Abstract

This chapter develops a typology of methods and measures for assessing the complexity of the built form at the scale of urban design. In particular, it extends quantitative methods from network science, ecosystems studies, fractal geometry, and information theory to the practice of neighborhood-scale urban design and the analysis of its qualitative human experience. Metrics at multiple scales are scattered throughout these bodies of literature and have useful applications in analyzing the built form that results from local planning and design processes. This chapter unpacks the connections between neighborhood-scale built form and measures of its complexity, and the typology developed here applies to empirical research of multiple neighborhood types and design standards. Finally, the typology includes several street network-analytic measures of urban form – emphasizing complexity in terms of density, permeability, resilience, and connectedness – applied in the subsequent empirical chapters.

## Introduction

This chapter examines measures of complexity appropriate for the spatial scale at which urban design tends to occur, namely the neighborhood and block scales. It arranges these measures into an analytical framework for assessing the complexity of the urban built form at the neighborhood scale. In particular, it extends quantitative methods from network science, ecology, fractal geometry, and information theory to the measurement of neighborhood-scale urban design and the analysis of its qualitative human experience. Metrics at multiple scales are scattered throughout these bodies of literature and have useful applications in analyzing the built form that results from local planning and design processes.

Rich linkages between complexity theory and urban design have been underexplored by researchers at the neighborhood and street scales – the scales of daily human experience and of the *practice* of urban design. The urban design literature frequently cites the value of “complexity” in neighborhood design, but these arguments often lack the theoretical formalism found in complex systems science. Nevertheless, following from Jane Jacobs (1961) and Christopher Alexander (1964; 1965), this body of scholarship argues that neighborhood complexity is essential to the life of the city and the function of its neighborhoods. Prominent design paradigms today, such as Smart Growth and the New Urbanism, frequently speak both directly and indirectly to complexity and notions of complex systems.

If neighborhood complexity is important, urban planners and designers require better tools to assess design outcomes and understand the built form. This chapter unpacks the connections between neighborhood-scale built form and measures of its complexity, and the typology developed here applies to empirical research of multiple neighborhood types and design standards. Finally, the typology includes several street network-analytic measures of urban form, applied in the subsequent empirical chapters. This chapter is organized as follows. First, it briefly reviews the lineage and meaning of “complexity” in urban design theory. Then it explores potential measures of the complexity of urban form at the scale of urban design projects, in four categories: temporal, visual, spatial, and structural. The structural measures are sub-divided into fractal and network measures. Finally, it organizes these various measures into a coherent typology.

## Background: Complexity in Urban Design

Urban designers often discuss physical urban form and design projects in terms of “complexity.” These discussions frequently borrow from the salient concepts of complex systems theory, but as discussed in chapter 3, often loosely, making it difficult to assess such claims and project outcomes. Nevertheless, various formulations of complexity have long been regarded as important in urban design, for several reasons. It can create more lively and enjoyable neighborhoods. It can improve urban resilience, robustness, connectivity, and access, playing into wider debates about sustainability and resource efficiency. Complexity can improve social equity and spatial distributional justice, and it can increase social contact, exchange, and adaptiveness. In this section I summarize this lineage of ideas about urban design and complexity – particularly through the notion that urban design influences the complexity of human habitats at the neighborhood scale and is closely tied to theories of livability. Then, in the subsequent section, I draw on complexity contextually and discuss several potential methods for measuring it.

### Urban Design, Livability, and Complexity

Urban design is the physical shaping of the public realm and borrows from both architecture and city planning. It includes deliberate top-down acts, informal bottom-up acts, and everything in between. The history of urban design has shifted through eras of classical formalism, romantic organicism, modernist simplifications, and post-Jane Jacobs gestures toward “organized complexity” (Jacobs 1961). In particular, urban design interfaces with complexity through notions of diversity, connectivity, resilience, and livability.

Livability has been defined in numerous ways and its meaning has evolved over time, but there is some common ground in the literature. Bosselmann (2008, p. 142) points out that “the original meaning of livability described conditions in neighborhoods where residents live relatively free from intrusions” but that the term has been progressively broadened to include sustainability, safety, comfort, available services, walkability, and transit. Macdonald (2005, p. 14) cites a modern vision of livable neighborhoods that create “lively, safe, and attractive streets, and [provide] public amenities such as parks, community centers, and schools.” Livability is in turn nested within even broader debates around urban sustainability and justice, as it is inextricably dependent on the city’s ability to meet all of its residents’ ongoing needs into the future. Several planning models – some competing, some complimentary – have taken up the mantle of livability in the U.S. today, including smart growth, the new urbanism, traditional neighborhood development, and transit-oriented development. Each promotes a compact urban form, walkability, and improved access to transit. Finally, issues of social justice cannot be ignored in the theorization of livability, as uneven distributions of power, capital, and privilege inevitably cloud the question of livability for whom and at the expense of whom (Evans 2002; Harvey 2010).

These definitions imply the importance of physical form and design for various aspects of livability. Indeed, livability is perhaps the key way in which planners engage with neighborhood form to address complexity. Three subcomponents of livability that particularly rely on complexity emerge from the literature. The first is visual complexity: an interrelation of qualities related to perceptible variety that makes public space lively, attractive, and enjoyable. The second is neighborhood completeness: a diverse mixture of amenities in close proximity. The third is connectivity of the circulation network. This body of literature argues that walkability and, in turn, livability rely on completeness and connectivity to be feasible and on visual complexity to be desirable. This has become a key goal of modern urban design, and will be discussed through the remained of section 4.3.

### The Neighborhood Scale

As indicated by the definitions above, livability and urban design interventions operate primarily at the scales of neighborhoods and blocks – usually with no more than a half mile radius (Boarnet and Crane 2001). Metrics for measuring the outcomes of urban design thus must consider the neighborhood scale. Neighborhoods are related to the concept of community, but also have a specific geographic, spatial nature (Larice and Macdonald 2007). They are ubiquitous around the world and play an important role in complex urban systems. Clarence Perry (2007, p. 55) said “an urban neighborhood should be regarded both as a unit of a larger whole and as a distinct entity in itself.” The concept of *neighborhood* has been theorized about by complexity scholars. For example, Portugali (2006) argues that cognitive conceptions of neighborhoods arise out of complex human systems via information compression, an idea based on the reduction of information in synergetics (Haken 2012).

According to Smith (2010, p. 137), “The spatial division of cities into districts or neighborhoods is one of the few universals of urban life from the earliest cities to the present.” Likewise, Mumford (1961, p. 193) points out that since the earliest days of cities, natural neighborhoods would form organically around important points like temples. Most pre-20th century neighborhoods were complete because walking was the most common mode of travel. Jackson (1985) describes such walking cities as dense and congested, with clear city/countryside distinctions and respectable locations nearest to the center of town, where accessibility was highest. Furthermore, “there were no neighborhoods exclusively given over to commercial, office, or residential functions… public buildings, hotels, churches, warehouses, shops, and homes were interspersed, or often located in the same structure” (ibid., p. 15). This draws together themes of completeness and connectivity.

Before the revolution in transportation technologies that culminated in the automobile, proximity was paramount and people lived near employment and retail. However, Allan Jacobs (1995, p. 311) suggests that in response to the dreadful living conditions of industrial-era cities and new enabling technologies, two major manifestos emerged to dominate 20th century neighborhood planning: Howard’s garden cities and modernism’s Charter of Athens. While the garden city movement largely respected the neighborhood, its legacy – sprawl – did not (Mumford 1961). Nor did Corbusier and the modernist planners, setting the stage for 20th century auto-dependency and single-use functional zoning (Hall 1996). This became the age of anti-complexity. Scott (1998) critiques the modernist design of Chandigarh and Brasilia by contrasting Le Corbusier’s top-down simplified, rational, polished, utopian cities with Jane Jacobs’s bottom-up, organically built, messy everyday urbanisms. He contends that the modernists confused geometric visual order for well-functioning sustainable social order in the built environment. Diversity, mixed uses, and complexity – grown naturally over time – make a community livable.

### Designing for Complexity

Scholars such as Jacobs (1961) argue that simplified single-purpose urban design destroys functional capacity and synergy. Over-simplified plans and interventions can cut into the living tissue of complex city systems, killing vital social processes. Healthy complex adaptive systems are resilient to perturbation, but their resilience and adaptability may be destroyed through too many simplifying interventions (Marshall 2012).

Yet every built environment has some deliberate design – especially in the public realm. Building facades are designed, roads are engineered, sidewalk widths are selected, and parks are laid out. Moroni (2010, p. 147) calls for rules that are simple, abstract, general, and purpose-independent to move away from a “flexible system of land-use planning” and toward “rules that enable society itself to be highly flexible.” To this end he suggests urban codes based on principles rather than details, contain few simple rules that remain for long periods of time, give minimal discretion to public officials, and leave flexibility for individual creativity and experimentation (Moroni 2015). Such codes already exist in urban design as form-based codes that aim to balance bottom-up flexibility with top-down predictability (Talen 2011). Marshall (2012, p. 203) similarly calls for a “system of planning” in which design and codes work together as a generative system that can give rise to a kind of emergent urbanism, with no guarantee that it will be optimal. However, development control can then be exercised to nudge what emerges toward the public interest (cf. Allen 2012). This is a middle ground between attempts to plan everything and attempts to plan nothing. Marshall suggests such a system would enable urban design and planning to deliver true functional complexity for neighborhoods. Jane Jacobs (1961) similarly argued that the role of planners is to generate diversity and supply what is lacking in a neighborhood.

According to this mainstream of scholarship, urban design and planning can foster diversity, connectedness, complexity, resilience, and robustness – elements of a healthy complex adaptive system. Yet an open question remains. Beyond these qualitative formulations of complexity in urban form and design, how might it be defined quantitatively – especially at the neighborhood and block scale? In the next section, I define complexity contextually and discuss several potential methods for measuring it.

## Measures of Complexity in Urban Form and Design

### Overview

Many complexity metrics at multiple scales, from metropolitan to neighborhood to building, are scattered throughout different bodies of literature. At a high level, Shiner and Davison (1999) examine complexity from the perspective of order and disorder. They present three types of complexity, and thus three types of measures. The first is positively correlated with disorder and includes algorithmic complexity and most measures of entropy. This type of complexity is highest when objects are scrambled up with the greatest variety and diversity. The second is a convex function of disorder, peaking at some midpoint between order and disorder. This type balances between variety and structure and conforms to traditional definitions of complex adaptive systems. The third takes complexity to be related to order or structure, and includes notions of self-organization and emergence in which structure emerges from previous disorder. Gershenson and Fernandez (2012) similarly argue that complexity is best described as a balance between self-organization and emergence, order and chaos. This chapter touches on all three categories as measures of aggregate complexity in urban design, depending on the context and character, but focuses on the second type: balancing between variety and structure.

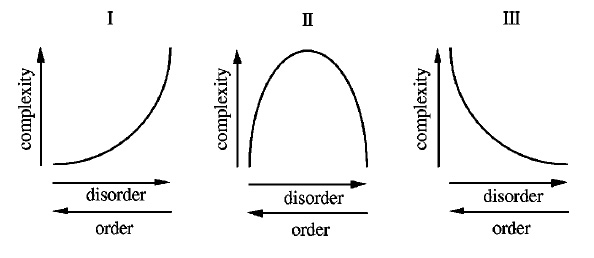


Figure 1. Three different notions of complexity. Type one complexity increases monotonically with disorder. Type two is convex, peaking at a midpoint between order and disorder. Type three decreases monotonically as disorder increases. Source: Shiner and Davison (1999)

Lloyd (2001) surveyed and categorized measures of complexity across numerous fields of inquiry. Bourdic et al. (2012) provide an overview of cross-scale spatial indicators and briefly touch on the neighborhood design assessment criteria of LEED-ND. Additional surveys and analyses of complexity indicators for ecosystems and cities have been produced by Parrott (2010) and Salat et al. (2010), mostly focusing on urban processes. If cities are complex systems, then there should be indicators of their complexity, which vary. Such an indicator of complexity could be a system-level state variable. However, just how to measure the complexity of a city system remains a topic of ongoing debate. Further, how does this sense – or preference – vary from person to person and culture to culture? We may have some intuitive sense of the complexity of a place simply by observing it or moving through it, but how can this be formalized?

On one hand, a neighborhood can be examined as an urban ecosystem – a human habitat – that is a complex dynamical system. State variables (such as population, density, employment, wealth, traffic volume, etc.) can be identified and calculated at various scales to describe the state of the system as it changes over time. The system’s dynamics can be explored and modeled with mathematical equations, statistical regressions, cellular automata, or agent-based models to describe the processes occurring in the system. Forrester (1969) was an early pioneer of applying systems dynamics analysis to cities, studying their stocks and flows. This use of differential equations and stock/flow modeling has been extended to cycles of urbanization and suburbanization (Orishimo 1987), and to the dynamics of parking (Cao and Menendez 2015). Such measures are less useful for the characterization and analysis of urban form.

On the other hand, a neighborhood can be examined as an output or product of human behavior and production. This focuses on the *physical form* of the neighborhood rather than its dynamical processes. Through co-evolution, as discussed in chapter 3, humans both shape their neighborhoods and are in turn shaped by them. The physical patterns that result are the urban form and can be examined in terms of network analyses, fractal structure, diversity (of various sorts), and information entropy. At a higher level of abstraction, neighborhood complex systems can be analyzed in terms of their resilience, robustness, and adaptiveness. How might the system’s dynamics respond to perturbation given its spatial patterns and structure in terms of connectedness and efficiency?

The following framework borrows, adapts, and reformulates relevant metrics to measure complexity at the neighborhood scale, touching on temporal measures but focusing on spatial and structural measures. In particular, it provides a quantitative framework that accounts for both traditional urban planning/design measures as well as more abstract measures arising from the complexity sciences. It is worth noting that this framework is not aimed at quantifying all aspects of “good” neighborhood design. Rather it intends to formalize and measure the indistinct notion of *complexity* as it applies to urban design. Qualities related to vitality, sustainability, sense of place, and other prominent qualities may overlap in some ways with the complexity metrics in this typology, but they are otherwise not the focus of this work.

### Temporal Measures of Urban Form

The first group of measures in this framework is temporal measures. Temporal measures describe time series data and in turn system dynamics. Such techniques include embedding the time series in state space, uncovering underlying attractors, estimating Lyapunov exponents (as discussed in chapter 2), and analyzing the system from an information theoretic perspective, such as Shannon entropy.

Nonlinear analysis techniques from the physical sciences, such as reconstructing attractors or estimating Lyapunovs, have not been found to be particularly effective in the ecology literature (Parrott 2010). Information theory, however, provides some measures of complexity that may be applied to urban design at the neighborhood scale. Shannon’s (1948) original theory of information entropy concerns the average amount of information contained in the revelation of a message or event. *Shannon entropy* indicates that the more types of things there are and the more equal each type’s proportional abundance is, the less predictable the type of any single object will be. This can be applied to abstract messages, time series, or spatial diversity, as discussed below. Entropy is lowest when the system is highly ordered and thus completely predictable. It is highest when the system’s disorder is highest. Such a type one measure thus emphasizes disorder rather than peaking at the point of (arguably maximum) complexity between order and disorder (Batty 2005, Yeh and Li 2001).

Derived from Shannon entropy, *mean information gain* assesses how much new information is gained from each subsequent datum in a time series (Proulx and Parrott 2008) and *fluctuation complexity* measures the amount of structure within a time series by evaluating the order of and relationship between values in the series: how likely is it we will observe some value *a* proximately after some other value *b*. Shannon entropy, mean information gain, and fluctuation complexity can be used to assess time series data arising from urban systems. However, more usefully, they might be abstracted and re-appropriated to evaluate the human experience of moving through the physical space that results from urban design.

### Visual Complexity of Urban Form

In a simplified, low-information urban landscape, little new information is gained by a pedestrian through the visual revelations of each passing step. In a highly complex urban environment (in terms of a type one measure), however, an individual will be bombarded with enormous amounts of new information as he or she moves through space. In these cases, space is the medium and the unfolding visual tableau is the message. This message could be discretized into arbitrary units such as meters, or into units relative to the specific urban landscape, such as street blocks or land parcels.

Clifton et al. (2008) discuss qualities of the urban form and human perceptions at multiple scales. For neighborhood and street scale urban design, perceptions of human scale are related to building heights and signage, perceptions of coherence are related to consistency of building heights, and sense of enclosure is related to building/element spacing and tree canopy. “Good” visual complexity tends to reach an optimum at some balance point between order and disorder, with “unity in variety” (Elsheshtawy 1997), implying a type two convex measure of complexity.

Ewing and Clemente (2013) performed a literature review that yielded 51 perceptual qualities of urban environments, eight of which were selected for further study because of their importance across the literature: imageability, enclosure, human scale, transparency, coherence, legibility, linkage, and visual complexity. These researchers related visual complexity to the number of perceptible differences a person is exposed to while moving through the city. They found that humans prefer to experience information at a comfortable rate – too little deprives the senses and too much overloads them.

Ewing and Clemente also found that good visual complexity depends on variety: types of buildings, design details, street furniture, signage, human activity, sunlight patterns, and the rich textural details of trees. Complexity is lost when design becomes too top-down, controlled, and predictable in modern large-scale master plans. Poor complexity exists when urban design elements are too few, are too similar and predictable, or are too disordered to be comprehensible (ibid.). In this formulation, complexity follows a type two convex function with a maximum value at some midpoint between order and disorder.

Based on their literature review, the researchers provide a field manual for measuring visual complexity. It is operationalized in five steps. First, count the number of buildings within the study area. Second, count basic and accent building colors. Third, record the presence of outdoor dining on each block as a binary value, present/not. Fourth, count the individual number of pieces of public art within the study area. Fifth, count the number of pedestrians in the study area. These measures of complexity are part of a larger toolkit for measuring urban design according to the eight perceptual qualities cited earlier (ibid.). Cavalcante et al. (2014) provide an alternate, statistical image processing measure of urban visual complexity.

Jacobs and Appleyard (1987) argue that buildings in varied arrangements add to visual complexity but interminable wide buildings – a hallmark of modernist design – detract from it. Jacobs (1995) argues that buildings need multiple varied surfaces for light to move constantly over to generate visual complexity. Macdonald (2005) explores how Vancouver generates visual complexity to put proverbial eyes on the street, with many entryways and interesting ground-level design.

Slow-moving pedestrians need a high level of complexity to hold their interest, but fast-moving motorists find that same environment chaotic. Dumbaugh and Li (2011) find that urban designs that balance vehicle speeds, visual complexity, and traffic conflicts can increase motorist awareness, decrease collisions, and improve pedestrian safety. Marshall (2012) contends that urban environments with perceptual richness are more interesting and enjoyable for humans, possibly because our species evolved in natural environments with a high degree of visual complexity. Thus, appropriate visual complexity is considered a key component of livability because it creates rich, enjoyable, safe environments for humans.

### Spatial Measures of Urban Form

The urban form that emerges from urban design is spatially embedded and can be characterized by various spatial measures of complexity. These measures assess the character of spatial patterns of the system at snapshots in time rather than looking at dynamics over time. Shannon entropy has been used to measure urban complexity (Batty 2005) and mean information gain has been used to measure ecosystem spatial complexity (Proulx and Parrott 2008). Yeh and Li (2001) used entropy to monitor and measure urban sprawl. Applying these information theoretic metrics to space usually entails assessing raster data for predictability.

Diversity, however, is the most common spatial measure of complexity in the urban design and planning literature. Diversity is important for several reasons. Social diversity can enhance learning, adaptation, and unexpected social mixing. Jane Jacobs (1961) praised diverse land uses for their ability to create synergies from complementary functions. Boarnet and Crane’s (2001) behavioral framework of the demand for travel fundamentally says that urban design influences the (time) cost of travel by placing origins and destinations in closer or further proximity to one another. Cervero and Kockelman (1997) also argue for land use diversity as a key feature shaping human travel behavior in urban environments.

Salat et al. (2010) identify three types of urban spatial diversity related to complexity: diversity among similar objects, diversity in spatial distribution, and diversity of scale. Diversity among similar objects refers to different characteristics of the same type of thing – for example, the “thing” might be humans and the characteristics might be income, race, employment, education, and so forth. It does however imply that *even distributions* are optimal in that they score the highest. This is a questionable reflection of complexity and a risky goal for central planning. Measures of dispersion and physical shape are also useful in characterizing the uniformity, randomness, or spatial complexity of ecosystems and could be applied to the built environment as well.

Wissen Hayek et al. (2015) use UrbanSim and measures of land use mix and density to evaluate the quality of the neighborhood-scale urban environment. The Simpson diversity index measures the diversity of objects in total across space, and is a common measure of land use entropy (i.e., land use mix) in the urban planning literature. This index is often called the Herfindahl-Hirschmann index in economics and the Probability of Interspecies Encounter in the ecology literature. This index is an *integral measure* that considers land use in a district as a whole, ignoring microscale structure and pattern (Song et al 2013):

In contrast, a *divisional measure* is sensitive to patterns *within* a district. This is a superior type of measure to consider questions of scale. The *dissimilarity index* measures how the land use mix within a district relates to the mix across the area as a whole – for two land use types, and for multiple (ibid.). Similar measures of dissimilarity are explored by Bordoloi et al. (2013). These spatial distributions of objects concern how equitably some set of desirable or undesirable objects is distributed across the city. For example, are all schools clustered in wealthy neighborhoods rather than being distributed evenly among all neighborhoods? Are waste treatment facilities clustered in poor neighborhoods rather than being distributed evenly among all neighborhoods? However, in a complex system, centers and clusters may form for inevitable or even “good” reasons. Agglomeration economies can cause job centers to cluster in certain areas. Ecosystem services of urban forests are highest when green spaces are concentrated and clustered rather than evenly distributed throughout urban development (Stott et al. 2015).

Diversity of scale addresses this specific issue further. Certain distributions within a complex system may be more efficient when they follow a power law or lognormal distribution rather than an even distribution. For example, it is not likely ideal for a neighborhood to have the same number of arterial roads, collector streets, and local streets. Rather, there might be a small number of large arterial roads, a medium number of mid-sized collector streets, and a large number of capillary local streets. Murcio et al. (2015) similarly use urban transfer entropy to examine multi-scale urban patterns and flows. Related to diversity, questions of scale and topological *structure* are addressed in the following section.

### Structural Measures of Urban Form: Fractal

Measures of topological structure assess the internal physical configuration of a system. They have been applied to cities and are perhaps the most useful measures of the complexity outcomes of urban design because they characterize that which is most dependent on the urban design process: physical structure and arrangement. Density itself might be a simple proxy for complexity as a greater number of things operating in the same area imply structure and connectivity. At the scale of urban design, these structural measures fall primarily into two categories: measures of fractal structure and network analysis.

Fractal structure refers to the “roughness” and self-similarity of some object, and how its detail relates to the scale at which it is observed. As discussed in chapter 2, fractals are self-similar, meaning that they have a similar structure at every scale. But in the real world, fractals are not perfect and do not exist at all spatial scales – from the infinitesimal to the infinite – as abstract mathematical fractals do. However, self-similarity of patterns and structure over multiple scales exist throughout nature. Batty (e.g., 2005) has long demonstrated how city structure and urban peripheries also are fractal.

Fractal structures tend to be distributed according to a power law. As briefly mentioned earlier, in a power law distribution, there are few large objects, a medium amount of medium sized objects, and very many small objects. Consider the earlier example of an urban street network. At the largest scale, the city has a few major arterial roads and boulevards that serve as the key arteries for system-wide traffic circulation. But if you zoom into this picture, a larger number of mid-sized collector roads appear, branching off from these few large arteries. As you zoom in further to a fine scale, a denser mesh of local streets appears, branching off from these collector roads. Similar fractal analyses have been applied to the distribution and scale of other urban structures such as buildings and land uses.

The fractal dimension, *D*, is a statistical measure of how a form’s complexity changes with regard to the scale at which it is measured:

6a0c8d27c7d09e77834816405c84e82a

eab68bd55aeb86348a8f02a0d8d990de

In these formulae, *N* is the number of new objects generated as scale transitions and *ε* is the scaling factor. This log-log ratio is similar to elasticities in economics. The fractal dimension of an object with one topological dimension refers to its space-filling characteristics that, through self-similarity, become a bit more than a one-dimensional line yet a bit less than a two-dimensional plane. Measures of fractal dimension include the Hausdorff dimension and the box-counting dimension (Shen 2002). For example, a Koch curve has a Hausdorff fractal dimension D = -log(4)/log(1/3) = 1.26.

The concept of fractal dimensions can also be applied to two dimensional surfaces, such as the surface of a city, the surface of a building, or the surface of elements of urban design. The fractal dimension is closely related to the qualities of visual complexity in urban design and public architecture, discussed earlier. While modernist architecture sought to erase complexity with simplified, segregated, sterile forms, both traditional architecture and today’s ideal tend to emphasize organic forms with rich detail *at multiple scales*. Salingaros (2001) argues that architecture and urban design must utilize fractal design to embrace the structure and organization of organic forms. The Eiffel Tower is an example of a built form that exhibits fractal structure. As Mandelbrot (1983, p. 131) puts it, “(well before Koch, Peano, and Sierpinski), the tower that Gustave Eiffel built in Paris deliberately incorporates the idea of a fractal curve full of branch points.”

### Structural Measures of Urban Form: Network

Beyond fractals, the second crucial lens with which to examine structure is network science. Accessibility is a useful measure of urban design and is related to network analysis. Accessibility concerns proximity, transportation mobility, and social interaction within the public sphere. Popular “walkability” tools – such as WalkScore and Walkonomics – and urban modeling tools such as UrbanSim/pandana use street networks to determine accessibility. Urban networks – considered here as primal, non-planar, weighted multidigraphs with self-loops – can be measured for their type 2 complexity based on their structure, particularly in terms of density, permeability, resilience, and connectedness. Extended definitions of and algorithms for the following measures can be found in Newman (2010) and Barthelemy (2011). The measures here are divided into metric measures and topological measures, but it is worth noting that in a planar graph, topological and metric properties are interrelated (Masucci et al. 2009).

*Metric structure* can be measured in terms of lengths (i.e., edge weights) and area and represents common transportation-design variables (e.g., Cervero and Kockelman 1997; Ewing and Cervero 2010). The *average street length*, the mean edge length in the undirected representation of the graph, serves as a linear proxy for block size and indicates how fine-grained or coarse-grained the network is. The *node density* is the number of nodes divided by the area covered by the network, and the *intersection density* is the node density of the set of nodes with more than one street emanating from them (thus excluding cul-de-sacs). The *edge density* is the linear sum of all edge lengths divided by the area, and the *street density* is the linear sum of all edges in the undirected representation of the graph divided by the area. These four density measures all provide further indication of how fine-grained the network is. Finally, the average circuity divides the sum of all edge lengths by the sum of the great-circle distances between the nodes incident to each edge (cf. Barthelemy 2011; Strano et al. 2012; Giacomin and Levinson 2015). This circuity measure is the average ratio between an edge length and the straight-line distance between the two nodes it links. All of these metrics are emergent properties of the complex organizing of the network (Masucci et al. 2009).

*Topological measures* of street network structure indicate the connectedness and robustness of the network, and how these values are distributed. The *average node degree*, or mean number of edges incident to each nodes, quantifies how well the nodes are connected, are average. Similarly but more concretely, the *average streets per intersection* measures the mean number of streets (i.e., edges in the undirected representation of the graph) that emanate from each intersection (i.e., node). This adapts the average node degree for physical form rather than circulation and flow. The distribution and proportion of number of streets per intersection characterizes the type, prevalence, and spatial distribution of intersection connectedness in the network.

The *eccentricity* of a node is the maximum of the shortest-path distances (weighted by length) between it and each other node in the network, and represents how far the node is from the node that is furthest from it. The *diameter* of a network is the maximum eccentricity of any node in the network and the *radius* of a network is the minimum eccentricity of any node in the network. The *center* of a network is the node or set of nodes whose eccentricity equals the radius and the *periphery* of a network is the node or set of nodes whose eccentricity equals the diameter. These distances measure network complexity in terms of size, structure, and shape.

*Connectivity* measures the minimum number of nodes or edges that must be removed from a connected graph to disconnect the network. This is a measure of robustness as complex networks with high connectivity provide more routing choices to agents and are more robust against failure. However, node and edge connectivity is less useful for approximately planar networks like street networks: *most* street networks will have connectivity of *1*, because the presence of a single cul-de-sac indicates that the removal of just one node or edge will disconnect the network. Rather, the *average node connectivity* of a network – the mean number of internally node-disjoint paths between each pair of nodes in G – more usefully represents the expected number of nodes that must be removed to disconnect a randomly selected pair of non-adjacent nodes (Beineke et al. 2002). Other measures of connectedness – such as intersection density, node degree distribution, and centrality/clustering (discussed below) – may capture the nature of a street network’s connectivity better than node or edge connectivity can. Networks with low connectivity may have multiple single points of failure, leaving the system particularly vulnerable. This can be seen in urban design through permeability and choke points: if circulation is forced through single points of failure, traffic jams ensue and circulation networks can fail. Connectivity has also been linked to street network pedestrian volume (Hajrashouliha and Yin 2015). Contrarily, Salingaros (2000) argues that grid networks do not connect cities, only giving the impression of doing so. He emphasizes eight characteristics – couplings, diversity, boundaries, forces, organization, hierarchy, interdependence, and decomposition – in his analysis of connectivity.

Network distances, degrees, and connectivity are significantly constrained by spatial embeddedness and approximate planarity (O’Sullivan 2014), so measures of clustering and centrality may better reveal topological structure and its distribution. The *clustering coefficient* of a node is the ratio of the number of edges between its neighbors to the maximum possible number of edges that could exist between these neighbors. The *weighted clustering coefficient* weights this ratio by edge length and the *average clustering coefficient* is the mean of the clustering coefficients of all the nodes in the network. These measure connectedness and complexity by how thoroughly the neighbors of some node are linked to each other. Jiang and Claramunt (2004) extend this coefficient to neighborhoods within an arbitrary distance, rather than just proximate, to make it more applicable to urban street networks.

Measures of centrality indicate the most important nodes in a network. *Betweenness centrality* assesses the importance of a node by evaluating the number of shortest paths that pass through it (Barthelemy 2004). The average betweenness centrality is the mean of betweenness centralities of all the nodes in the network (Barthelemy 2011). Betweenness centrality can also be calculated for weighted networks (Barrat et al. 2004). Barthelemy et al. (2013) uses betweenness centrality to identify top-down interventions versus bottom-up self-organization and evolution of the urban fabric in Paris. *Closeness centrality* represents, for each node, the reciprocal of the sum of the distance from this node to all others in the network (optionally weighted by length): that is, nodes rank as more central if they are on average closer to all other nodes. The *average closeness centrality* is the mean of the closeness centralities of all the nodes in the network. Finally, *pagerank* (the algorithm Google uses to rank web pages) is a variant of network centrality, namely eigenvector centrality (Brin and Page 2012). Pagerank ranks nodes based on the structure of incoming links and the rank of the source node, and may also be applied to street networks (Agyzkov et al. 2012, Chin and Wen 2015).

Measures of centrality are typically used in combination to assess street networks. Porta et al. (2006a; 2006b) demonstrate a multiple centrality assessment methodology for analyzing urban street networks and identify signatures and differences between planned and self-organized cities. Crucitti et al. (2006) examine closeness, betweenness, and information as measures of urban network centrality. The Urban Network Analysis Toolbox (Sevtsuk and Mekonnen 2012) analyzes betweenness centrality, closeness centrality, and accessibility in urban street networks.

Finally, it is worth mentioning space syntax theory and dual graphs. The street networks discussed so far are *primal*: the graphs represent intersections as nodes and street segments as edges. In contrast, a *dual graph* inverts this network topology: a city’s streets are represented as nodes and the intersections are represented as edges. Such a representation seems a bit odd, but provides certain advantages in analyzing the network topology (Crucitti et al. 2006). Dual graphs form the foundation of space syntax, another method of analyzing urban networks and configuration. Space syntax analyzes axial street lines and measures the depth from some network edge to others (Hillier et al. 1976; cf. Ratti 2004). Marcus and Legeby (2012) use space syntax to measure social capital in neighborhoods, through an explicit urban complexity lens. Jiang and Claramunt (2002) integrate an adapted space syntax – compensating for difficulties with axial lines – into computational GIS. Space syntax has formed the basis of many other adapted approaches to analytical urban design (e.g., Karimi 2012). This present study, however, focuses almost entirely on primal graphs because they retain all the geographic, spatial, metric information essential to urban form and design, that space syntax must disregard in its dual graphs.

## Preliminary Typology of Complexity Measures

All of these methods of assessing the complexity of urban design, primarily at the neighborhood scale, can be fit together into a preliminary typology. Here the measures are grouped into four types: temporal, spatial, visual, and structural. The structural measures are subdivided into fractal and network measures. Spatial, visual, and structural measures seem to be the most promising for measuring physical complexity at the scale of urban design.

|  |  |  |
| --- | --- | --- |
| Category | Measure of complexity | Description |
| Temporal | Embedding time series | Examine variables in state space to reveal possible deep structure and patterns in data |
| Temporal, Spatial | Shannon entropy | How unpredictable a sequence is, based on number of types and proportional abundance |
| Temporal, Spatial | Mean information gain | How much new information is gained from each subsequent datum |
| Temporal | Fluctuation complexity | Amount of structure within a time series |
| Temporal, Spatial | Urban Transfer Entropy | Analytic tool for examining multi-scale urban patterns and flows |
| Visual | Ewing and Clemente field guide | Set of methods for assessing the physical, visual complexity of the streetscape |
| Visual | Cavalcante streetscape measure | Image processing method to assess visual complexity on contrast and spatial frequency |
| Spatial | Simpson diversity index | Assesses land use mix: how homogenous or heterogenous is the area of analysis? |
| Spatial | Dissimilarity index | How does the land use mix within a subarea relate to the mix across the entire area? |
| Fractal | Hausdorff fractal dimension | How a form’s complexity changes with regard to the scale at which it is measured |
| Fractal | Box-counting fractal dimension | How a form’s complexity changes with regard to the scale at which it is measured |
| Spatial,  Network | Destination accessibility | A function of land use entropy, amenity distribution, and network structure |
| Network | Average streets per intersection | How well connected and permeable the physical form of the street network is, on average |
| Network | Proportion of streets per intersection | Characterizes the type, prevalence, and spatial distribution of intersection connectedness |
| Network | Average street length | How long the average block is between intersections; proxy for block size |
| Network | Node/intersection, edge/street density | How fine- or coarse-grained the street network is |
| Network | Average circuity | How similar network distances are to straight-line distances |
| Network | Diameter/periphery, radius/center | Measure network complexity in terms of max/min size, structure, and shape |
| Network | Node/edge connectivity | What is the minimum number of nodes/edges that must be removed to disconnect network? |
| Network | Average node connectivity | Average number of nodes that must be removed to disconnect some pair of non-adjacent nodes |
| Network | Clustering coefficient | Extent to which the neighbors of some node are linked to each other |
| Network | Average clustering coefficient | Mean of the clustering coefficients for all nodes |
| Network | Betweenness centrality | The importance of a node in in terms of how many shortest paths use that node |
| Network | Average betweenness centrality | Mean of the betweenness centralities for all nodes |
| Network | Closeness centrality | Nodes rank as more central if they are on average closer to all other nodes |
| Network | Average closeness centrality | Mean of the closeness centralities for all nodes |
| Network | Pagerank | Ranking of node importance based on structure of incoming links |
| Network | Multiple centrality assessment | Uses primal, metric graphs to examine multiple indices of centrality |
| Network | Space syntax | Uses dual, topological graphs to examine closeness centrality of a named street |

Table 4.1 Typology of measures of the complexity of urban form/design.

## Discussion

Practitioners and theorists have expounded on complexity’s value long before and long after the days of Jane Jacobs. Complexity underlies urban resilience and sustainability planning (Jabareen 2013; Mattsson and Jenelius 2015). Path dependence, hysteresis, and historical accidents all arise from complex systems and drastically affect the trajectory of urban form (Siodla 2015): these features are both products of urban design and constraints to urban design. More complex urban environments are more resilient and robust. They provide greater opportunities for social encounter, mixing, and adaptation through social learning. Complexity entails greater connectivity, diversity, and variety – all of which can improve social justice and sustainability. Today, prominent urban design movements such as the New Urbanism and Smart Growth openly embrace the notion of complexity.

This chapter sought to refine what “complexity” means in this context and in turn provide ways to measure it. Of the three types of complexity (in terms of order/disorder) presented at the beginning of this chapter, type one seems most appropriate for measuring the complexity of *difference*: how scrambled up land uses and socioeconomic traits are. The second type seems most appropriate for *organized complexity*, where a balance between chaos and order is desirable, such as in visual complexity and structural complexity. Likewise, the third measure is most useful when looking for *ordered* elements of urban design, particularly when some have self-organized from an original disordered state.

The preliminary typology presented in this chapter draws from different scientific disciplines to offer different measures of complexity that apply to urban form and particularly to urban design’s scale of intervention. In particular, the measures of network structure characterize the complexity of the circulation network in terms of density, permeability, resilience, and connectedness. These attributes influence the way an urban system’s physical links can structure complex interactions, connections, and dynamics. The analytical framework developed here is generalizable to empirical research of multiple neighborhood types and design standards. In particular, network-analytic measures in this typology are applied empirically in the next two chapters. Chapter 5 presents a new toolkit for acquiring, constructing, analyzing, and visualizing urban street networks and demonstrates it with a case study. Chapter 6 conducts a large empirical study of street networks at multiple scales using the metrics introduced in chapter 4 and the toolkit introduced in chapter 5.

# Acquiring, Constructing, Analyzing, and Visualizing Street Networks

## Abstract

Scholars have studied street networks in various ways. However, there are some limitations to the current urban planning/street network analysis literature. To address these challenges, this study presents a new tool to make the collection of data and creation and analysis of street networks simple, consistent, and automatable. OSMnx is a new Python package that downloads administrative boundary shapes and street networks from OpenStreetMap. OSMnx contributes five significant new capabilities for researchers and city planners: first, the automatic downloading of administrative place boundaries and shapefiles; second, the tailored and automated downloading and constructing of street networks from OpenStreetMap; third, the automated correction and simplification of network topology; fourth, the ability to save street networks to disk as shapefiles, GraphML, or SVG files; and fifth, the ability to analyze street networks, calculate routes, project and visualize the networks, and calculate network metrics and statistics. These metrics and statistics include both those common in urban design and transportation studies, and measures of the structure and topology of the network.

## Introduction

Urban scholars and planners have studied street networks in various ways. Some studies have focused on the urban form, others on transportation and flow, and others on the topology, complexity, and resilience of street networks. However, the current literature suffers from some shortcomings. First, the sample sizes in cross-sectional studies tend to be quite small. Second, studies usually simplify the representation of the street network to a planar or undirected primal graph for tractability. Third, the dozens of decisions that go into analysis – such as spatial extents, topological simplification and correction, definitions of nodes and edges, etc. – are often ad hoc or only partly reported, make reproducibility challenging. Fourth, the current landscape of tools and methods offers no ideal technique that balances usability, customizability, reproducibility, and scalability in acquiring, constructing, and analyzing network data.

To address these challenges, this study presents OSMnx, a new Python package that downloads administrative boundary shapes and street networks from OpenStreetMap. OSMnx contributes five significant new capabilities for researchers and city planners: first, the automatic downloading of administrative place boundaries and shapefiles; second, the tailored and automated downloading and constructing of street networks from OpenStreetMap; third, the automated correction and simplification of network topology; fourth, the ability to save street networks to disk as shapefiles, GraphML, or SVG files; and fifth, the ability to analyze street networks, calculate routes, project and visualize the networks, and calculate network metrics and statistics. These metrics and statistics include both those common in urban design and transportation studies, as well as metrics that measure the structure and topology of the network.

This chapter is organized as follows. First it discusses the background of networks, street network analysis and representation, and the current landscape of tools for this type of research – including their shortcomings. Next it presents OSMnx and discusses its functionality. Then it presents a simple case study, using OSMnx to construct and analyze street networks in Portland, Oregon. Finally, it ends with a brief conclusion.

## Background

### Street network analysis

Street networks are spatial networks (discussed in chapter 3). Scholars have studied street networks in many ways.

However there have been very large-scale few studies of the complexity of large sets of street networks.

### Representation of street networks

Urban street networks are commonly represented using either a primal or a dual approach. A primal network represents streets as edges and intersections as nodes, much as we would typically think of a network. However, a dual network inverts this topology, representing streets (i.e., each named street is a single entity) as nodes and intersections as edge. Certain network metrics, such as those based on a street’s connectivity or centrality, are easier to calculate using a dual approach. Space syntax theory represents individual named streets as single entities rather than multiple edges. However, all the spatial and geographic information of the street (such as its length, shape, circuity, width, etc.) are lost in a dual network. A primal network, in contrast, can faithfully represent all the spatial characteristics of a street. Primal may be a better approach for analyzing spatial networks when geography matters, because the physical space underlying the network contains relevant information that cannot exist in the network’s topology alone.

Typically, street networks are assembled into some sort of graph-theoretic object from GIS data. One way to construct a street network is to take the lines of all the streets in some study area and use a GIS to split them wherever they cross. These split segments become edges and the splitting points become nodes. However, this method assumes a planar graph: bridges and tunnels become splitting points (and thus nodes) even if the streets do not actually intersect in the real world. Unless the street network is truly planar, a planar representation is thus less-than-ideal simplification: such a street network would yield inaccurate analyses and metrics as the lengths of edges would be underestimated and the number of nodes would be overestimated.

The urban street network analysis literature tends to suffer from three problems. The first, as just discussed, is that the networks under analysis tend to be simplified to a planar network. This may reasonably represent a street network in a European medieval city center, but poorly represents the street network in a city like Oakland, California, with several grade-separated expressway, bridges, and tunnels in a truly non-planar street network. The second problem is sample size. Most cross-sectional studies in this body of literature tend to analyze fairly small sets of networks (at either the city or neighborhood scale) for tractability. Indeed, it can be extremely difficult and time-consuming to acquire and assemble large numbers of street networks from data sources and repositories spread across various governmental entities. The third problem is replicability. These studies tend to gloss over the precise details of how their street networks were constructed, yet numerous unreported decisions had to be made in the process. For example, some studies study entire cities, out to the urban periphery, but do not explain how the periphery was defined (e.g., Strano et al. 2013). Further, what is an edge in the street network? Some studies do not report if the network is directed or not. What is a node in the street network? Is it where at least two different named streets come together? Does it denote a junction of routes? Different studies make different decisions on these questions. When studies are performed ad hoc or with custom tools, many such decisions go unreported, and replicability becomes nigh impossible.

### Current tool landscape

There are several tools to study street networks, and many of them address these problems, but only in part. ArcGIS provides a Network Analyst extension, for which Sevtsuk and Mekonnen (2012) developed the Urban Network Analysis Toolkit plug-in. QGIS, an open-source alternative to ArcGIS, also provides some limited graph and network analysis capabilities through built-in tools and plug-ins. These GIS tools provide only limited network analysis abilities, such as shortest path calculations. Gephi is a popular network analysis software program, but does not natively provide the GIS functionality essential to study spatial networks. Pandana is a Python package that enables accessibility queries over a spatial network, but does not support other graph-theoretic network analyses (Foti 2014). Finally, NetworkX is a Python package for network analysis, supported by the Center for Nonlinear Studies at Los Alamos National Laboratory and the Complexity Sciences Center at UC Davis. NetworkX is free, open-source, and able to analyze networks with millions of nodes and edges (Schult and Swart 2008; Hagberg and Conway 2010).

Street network data comes from many sources, including city, state, and national data repositories, and typically in ESRI shapefile format. In the United States, the census bureau provides free TIGER/Line (Topologically Integrated Geographic Encoding and Referencing) shapefiles of geographic data such as cities, census tracts, roads, buildings, and certain natural features. However, TIGER/Line suffers from some inaccuracies. Further, the roads shapefiles contain MTFCC codes to identify route types, but they are very coarse-grained (e.g., classifying parking lots as alleys) and topologically depict bollarded intersections as through-streets (which is obviously a problem for routing).

OpenStreetMap has emerged in recent years as a major player both for mapping and for spatial data. OpenStreetMap is a collaborative mapping project that provides a free and publicly editable map of the world. Inspired by Wikipedia’s mass-collaboration model, the project started in 2004 and has grown to over two million users today. Its data quality is generally quite high (for instance, OpenStreetMap data is now the default in Garmin GPS devices), but the quality and coverage varies worldwide. In the United States, OpenStreetMap imported the 2005 TIGER/Line roads in 2007 as a foundational data source. Since, then numerous edits and corrections have been made. But more importantly, many additions have been made beyond what TIGER/Line captures, including pedestrian paths through parks, passageways between buildings, bike lanes and routes, and richer attribute data describing the characteristics of features, such as finer-grained codes for classifying arterial roads, collector streets, residential streets, alleys, parking lots, etc.

There are several ways of acquiring street network data from OpenStreetMap. First, OpenStreetMap provides an API, called Overpass, which can be queried programmatically to retrieve any data in the database: streets and otherwise. However, its usage and syntax are notoriously challenging and several services have thus sprung up to simplify the process. Mapzen extracts chunks of OpenStreetMap data constrained to bounding boxes around 200 metropolitan areas worldwide. They also provide custom extracts, which can take up to an hour to run. Mapzen works well for simple bounding boxes around popular cities, but otherwise does not provide an easily scalable or customizable solution. Geofabrik similarly provides data extracts, generally at the national or sub-national scale, but provides shapefiles as a paid service.

Finally, GISF2E is a tool (compatible with ArcGIS and an outdated version of Python) that can convert shapefiles (such as Mapzen or Geofabrik extracts) into graph-theoretic network data sets (Karduni 2016), and the creators provide processed shapefiles for several cities online but with several limitations. First, while GIS2FE shapefiles’ roads have a flag denoting one-way streets, it discards *to* and *from* nodes, so it is unclear which direction the one-way goes. Regarding the question posed earlier about what is a node in a street network, GIS2FE treats it inconsistently. Sometimes a right-angle is considered a node (i.e., an intersection at which two perpendicular named streets dead-end). Other times a right-angle is not considered a node, (i.e., a single named street turns 90 degrees). Topologically and spatially, these two cases are identical. But they are treated differently based on arbitrary break points between OpenStreetMap IDs or line digitization. OpenStreetMap IDs are sometimes 1-to-1 with a named street, other times a named street might have multiple OpenStreetMap ID segments. Further, some streets have arbitrary nodes in the middle of them because the OpenStreetMap ID is different on either side.

### Summary

The empirical literature on street network analysis is growing ever richer, but suffers from some challenges. First, sample sizes tend to be fairly small due to data availability, gathering, and processing constraints. Second, reproducibility is difficult when the dozens of decisions that go into analysis – such as spatial extents, topological simplification and correction, definitions of nodes and edges, etc. – are ad hoc or only partly reported. Third, and related to the first two, studies frequently oversimplify to planar or undirected primal graphs for tractability, or use dual graphs despite the loss of geographic and metric information. Fourth, the current landscape of tools and methods offers no ideal technique that balances usability, customizability, reproducibility, and scalability in acquiring, constructing, and analyzing network data.

## OSMnx: Functionality and comparison to existing tools

To address these challenges, the primary contribution of this dissertation is the creation of a tool to make the collection of data and creation and analysis of street networks simple, consistent, and automatable. OSMnx is a new Python package that downloads administrative boundary shapes and street networks from OpenStreetMap. It allows users to easily construct, project, visualize, and analyze non-planar complex street networks consistently in Python with NetworkX. Users can construct a city’s or neighborhood’s walking, driving, or biking network with a single line of Python code.

OSMnx contributes five significant new capabilities for researchers and city planners: first, the automatic downloading of administrative place boundaries and shapefiles; second, the tailored and automated downloading and constructing of street networks from OpenStreetMap; third, the automated correction and simplification of network topology; fourth, the ability to save street networks to disk as shapefiles, GraphML, or SVG files; and fifth, the ability to analyze street networks, calculate routes, visualize the networks, and calculate network metrics and statistics. These metrics and statistics include both those common in urban design and transportation studies, as well as metrics that measure the complexity of the network. The following subsections discuss each of these contributions in order.

### Acquiring administrative place boundaries

To acquire administrative boundary GIS data, one typically must track down shapefiles online and download them. However, bulk or automated acquisition and analysis (such as that required to analyze hundreds or thousands of separate geographies) requires clicking through numerous web pages to download shapefiles one at a time. With OSMnx, one can download place shapes from OpenStreetMap in a single line of Python code, and project them to UTM in one more line of code (all of the built-in projection in OSMnx calculates UTM zones automatically based on the centroid of the geometry). One can just as easily acquire polygons for other place types, such as neighborhoods, boroughs, counties, states, or nations – any place geometry available in OpenStreetMap. Or, one can pass multiple places into a single query to construct a single shapefile with multiple features from their geometries. This can also be done with cities, states, countries or any other geographic entities, and the results can be saved as a shapefile to a hard drive (Figure 5.1).

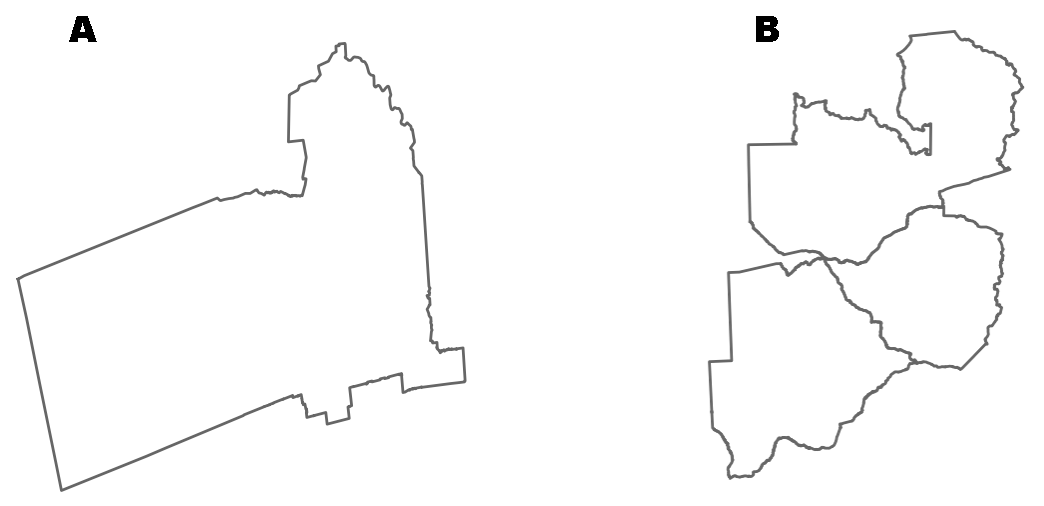


Figure 5.1 Administrative boundary vector geometries retrieved for A) Berkeley, California, and B) Zambia, Zimbabwe, and Botswana.

### Download and construct street networks

The primary contribution of OSMnx is the downloading and construction of street networks. To acquire street network GIS data, one must typically track down TIGER/Line roads from the US census bureau, or individual data sets from other countries or their cities. However, this becomes preventively burdensome for large numbers of separate street networks as it does not entail bulk, automated analysis. Further, it ignores informal paths and pedestrian circulation that TIGER/Line lacks? Finally, TIGER/Line provides no street network data for outside the United States. In contrast, OSMnx handles all of these use cases.

OSMnx lets one download street network data and build topologically-corrected street networks, project and plot the networks, and save the street network as SVGs, GraphML files, or shapefiles for later use. The street networks are directed graphs and preserve one-way directionality. One can download a street network by providing OSMnx any of the following queries:

* a bounding box
* a latitude-longitude point plus a distance in meters (either a distance along the network or a distance in each cardinal direction from the point)
* an address plus a distance in meters (also, either a distance along the network or a distance in each cardinal direction from the point)
* a polygon of the desired street network’s boundaries
* a place name or list of place names

One can also specify several different network types:

* ‘drive’ - get drivable public streets (but not service roads)
* ‘drive\_service’ - get drivable public streets, including service roads
* ‘walk’ - get all streets and paths that pedestrians can use (this network type ignores one-way directionality)
* ‘bike’ - get all streets and paths that cyclists can use
* ‘all’ - download all (non-private) OpenStreetMap streets and paths
* ‘all\_private’ - download all OpenStreetMap streets and paths, including private-access

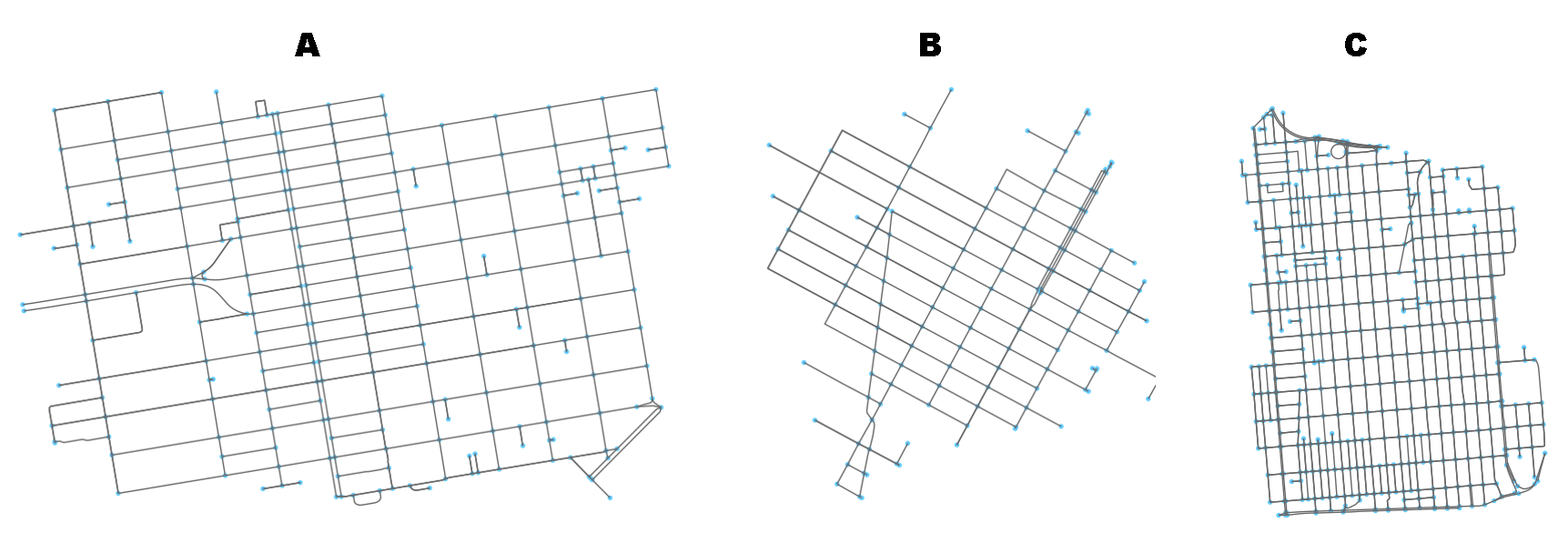


Figure 5.2 Street networks created by A) bounding box, B) address and network distance, and C) neighborhood polygon

The functionality to acquire street networks by place name or by polygon is particularly useful for researchers and planners. When passed a place name, OSMnx geocodes the name using OpenStreetMap’s Nominatim API and constructs a polygon from its administrative boundaries. It then buffers this polygon by 500 meters and downloads the street network data within its geometry from OpenStreetMap’s Overpass API. Next it constructs a street network from this data, corrects the topology, calculates accurate degrees and intersection types per intersection (this ensures that intersections are not considered cul-de-sacs simply because an incident edge connects to a node outside the desired polygon), then truncates the network to the original, desired polygon. One can just as easily request a street network within a borough, county, state, or other geographic entity. One can also pass a list of places (such as several neighboring cities) to create a unified street network within the union of their geometries.

The geometric operations take advantage of an R-tree spatial index to quickly identify nodes that lay within/outside of the polygon(s). An R-tree represents individual objects and their bounding boxes (the “R” is for “rectangle”) as the lowest level of the spatial index. It then aggregates nearby objects and represents them with their aggregate bounding box in the next higher level of the index. At yet higher levels, the R-tree aggregates bounding boxes and represents them by their bounding box, iteratively, until everything is nested into one top-level bounding box. To search, the R-tree takes a query box and, starting at the top level, sees which (if any) bounding boxes intersect it. It then expands each intersecting bounding box and sees which of the child bounding boxes inside it intersect the query box. This proceeds recursively until all intersecting boxes are searched down to the lowest level, and returns the matching objects from the lowest level. However, an R-tree provides no speed-up when the features’ bounding boxes are (approximately) identical, because it identifies (approximately) *every* point as a possible match as the bounding box of the polygon intersects every nested rectangle inside the index. This is a limitation of R-trees themselves. Instead, OSMnx subdivides the polygons into smaller sub-polygons with concomitantly smaller minimum bounding boxes. Then it iterates through these small sub-polygons to quickly identify which points lie within each, taking full advantage of the R-tree index. This drops the spatial processing time of, say, metropolitan Los Angeles from approximately an hour to approximately a few seconds.

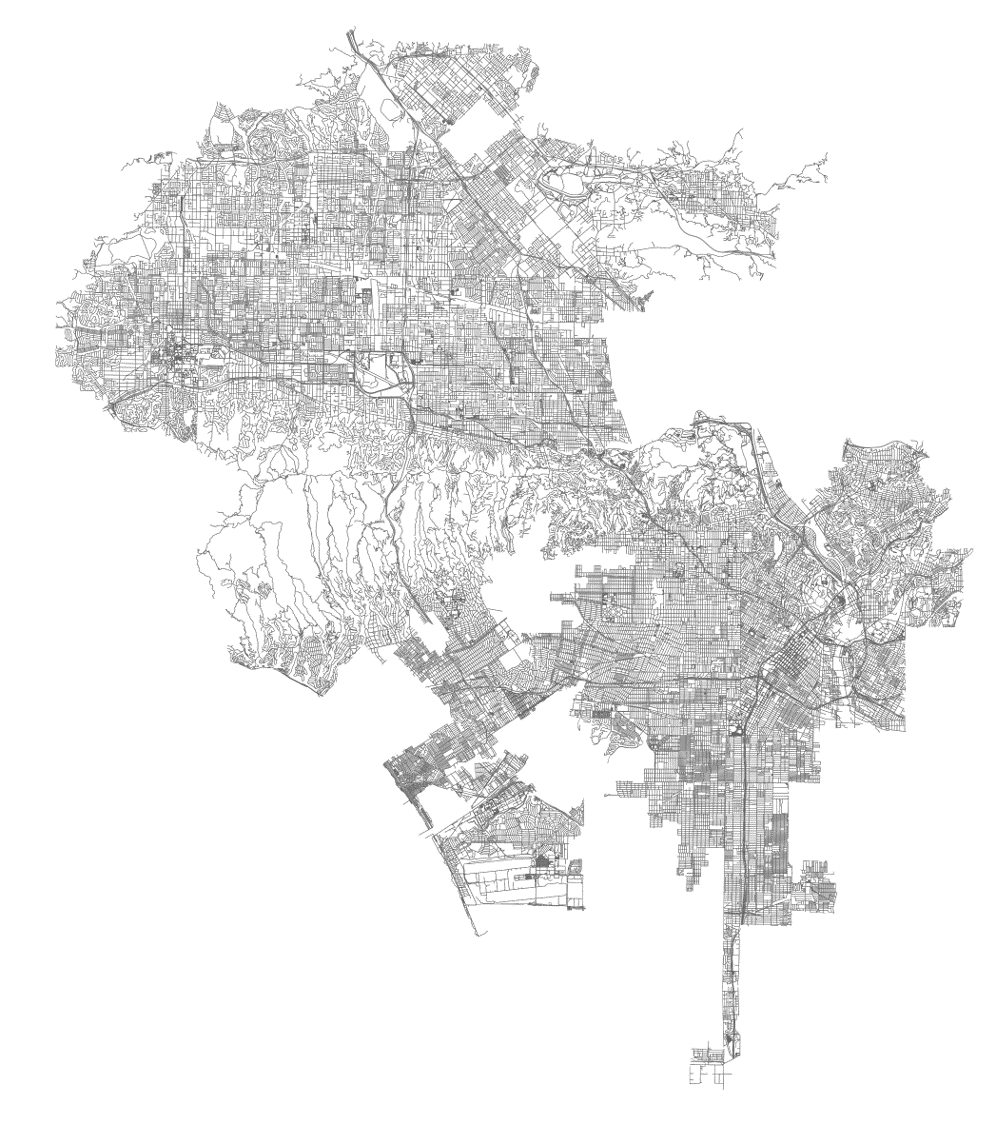


Figure 5.3 The drivable street network for municipal Los Angeles, created by simply passing the query phrase "Los Angeles, CA, USA" into OSMnx.

Also relevant to planner scholars and practitioners, OSMnx enables the acquisition of street networks around the world. In general, US street network data is fairly easy to come by thanks to TIGER/Line shapefiles. OSMnx makes it *easier* by making it available with a single line of code, and *better* by supplementing it with all the additional data (both attributes and non-road routes) from OpenStreetMap. However, one can just as easily acquire street networks from anywhere in the world - places where such data might otherwise be inconsistent, difficult, or impossible to come by (Figure 5.4).



Figure 5.4 Street networks for A) Modena, Italy, B) Belgrade, Serbia, C) central Maputo, Mozambique, and D) central Tunis, Tunisia.

### Correct and simplify network topology

Topological correction and simplification is performed by OSMnx automatically under the hood, but it is illuminating to break it out to see how it works. Simplification is essential for a correct topology because OpenStreetMap nodes can be inconsistent: they include intersections, but they also include all the points along a single street segment where the street curves. The latter are not nodes in the graph-theoretic sense, so we remove them algorithmically and consolidate the set of edges between “true” network nodes (i.e., intersections) into a single unified edge. These edges unified edges between intersections retain the full spatial geometry of the consolidated sub-edges and the relevant attributes, such as the length of the street segment. OSMnx provides different simplification modes to provide researchers fine-grained control to define nodes rigorously. In *strict* simplification mode, a node is either:

1. where an edge dead-ends (i.e., the dead end of a cul-de-sac), or
2. the endpoint from which an edge self-loops, or
3. the intersection between multiple streets where at least one of the streets continues *through* the intersection (i.e., if two streets dead-end at the same point, creating an elbow, the point is not considered a node)

In *non-strict* mode, conditions *1* and *2* remain the same, but *3* is changed to permit nodes at the intersection of two-streets, even if both streets dead-end there, as long as the streets have different OpenStreetMap IDs. In either mode, a node is retained if there is a point where a single street changes from one-way to two-way. The process of simplification is illustrated in figure 5.5. When we first download and assemble the street network from OpenStreetMap, appears as depicted in figure 5.5A. For one-way streets, directed edges are added from the origin node to the destination node. For two-way streets, directed edges are added in both directions between nodes. We want to simplify this network to only retain the nodes that represent the junction of multiple streets. OSMnx does this automatically in strict mode, unless told to do otherwise. First, it identifies all non-intersection nodes (i.e., all those that simplify form an expansion graph), as depicted in figure 5.5B. Then it removes them, but faithfully maintains the spatial geometry and attributes of the street segment between the true intersection nodes. In figure 5.5C, all the non-intersection nodes have been removed, all the true intersections (junctions of multiple streets) remain in blue, and self-loop nodes are in purple. In strict mode, OSMnx considered two-way intersections to be topologically identical to a single street that bends around a curve. Conversely, if one wishes to retain these intersections when the incident edges have different OSM IDs, he or she may use non-strict mode, as depicted in figure 5.5D.

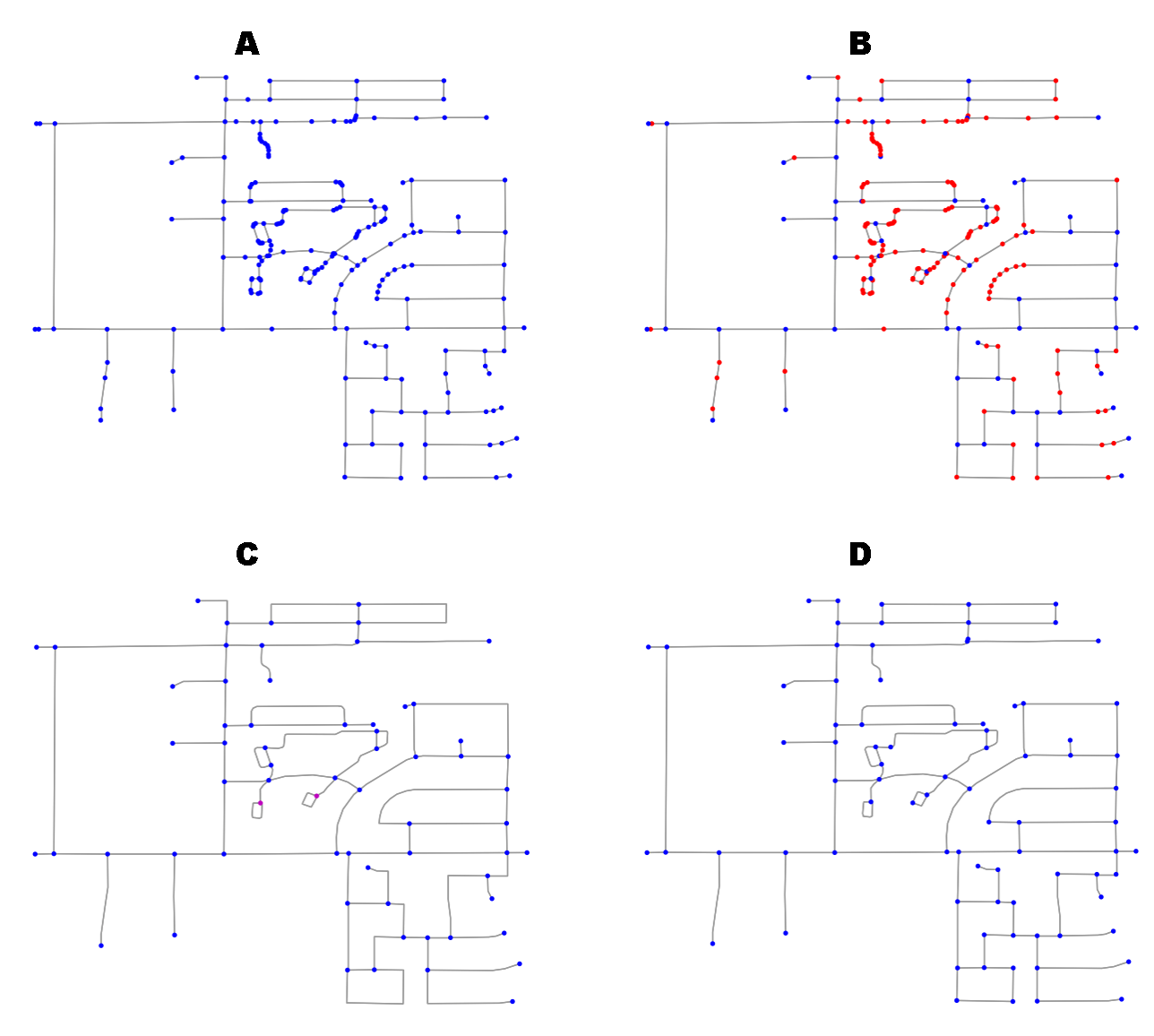


Figure 5.5 A) the original graph, B) non-graph-theoretic nodes highlighted in red and true intersections in blue, C) strictly simplified network, with self-loops noted in magenta, D) non-strictly simplified network.

### Save street networks to disk

OSMnx can save the street network to disk as a GraphML file (an open, standard file format for representing graphs on disk) to work with later in Gephi or NetworkX. Or it can save the network as ESRI shapefiles of nodes and edges to work with in any GIS. When saving as shapefiles, the network is simplified to an undirected representation; however, one-way directionality and origin/destination nodes are preserved as edge attributes for GIS routing applications, unlike GISF2E. OSMnx can also save street networks as scalable vector graphics (SVG) files for design work in Adobe Illustrator (Figure 5.6).

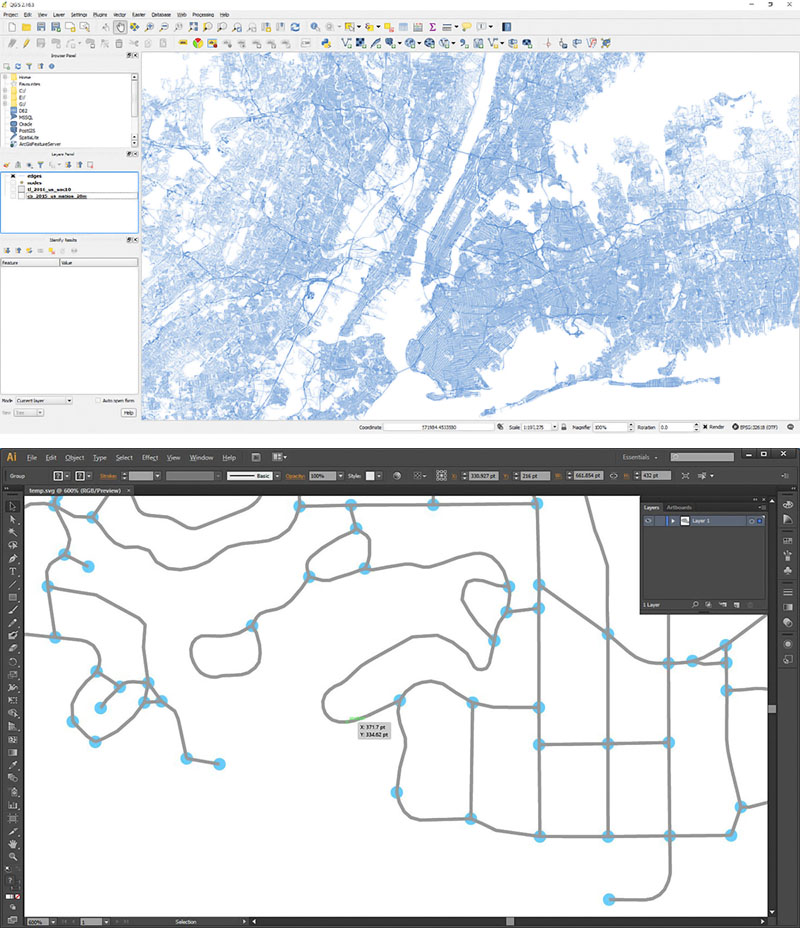


Figure 5.6 Street network for metropolitan New York from OSMnx saved and loaded in QGIS as an ESRI shapefile (above) and in Adobe Illustrator as SVG (below).

### Analyze street networks

OSMnx easily analyzes networks and calculates network statistics, including spatial metrics based on geographic area or weighted by distance (Table 5.1). With a single command, OSMnx calculates the *nodes*’ average neighborhood degrees (weighted and unweighted, betweenness centralities, closeness centralities, degree centralities, clustering coefficients (weighted and unweighted), pageranks and the *network*’s ), intersection count, intersection density, average betweenness centrality, average closeness centrality, average degree centrality, eccentricity, diameter, radius, center, periphery, node connectivity, average node connectivity, edge connectivity, average circuity, linear edge density per square kilometer, total edge length, average edge length, average degree, number of edges, number of nodes, node density per square kilometer, maximum and minimum pagerank values and corresponding nodes, the proportion of edges that self-loop, linear street density per square kilometer, total street length, average street length, number of street segments, average number of street segments emanating from each intersection, and the counts and proportions of node types.

The counts and proportions of streets per node improves on metrics used in other studies. Most studies simply use the node degree (either directed or undirected) to count how many edges are incident to the node. This has a couple of problems. First, it ignores the fact that nodes only have incident edges that connect to other nodes *within the spatial boundaries* requested. Thus, a 4-way intersection near the periphery may only have 3 incident edges in the network, because the fourth links a node outside the bounds and was therefore truncated. Second, bi-directional self-loops in a directed graph would count as *four* physical street segments connected to the intersection. This study instead uses a new algorithm to correctly count physical streets per intersection, taking into account one-way and two-way streets, parallel edges, self-loops, and intersections with streets that do not appear in the graph because they link to a node outside the bounds.

|  |  |
| --- | --- |
| Metric/statistic | Definition |
| *n* | number of nodes in the graph |
| *m* | number of edges in the graph |
| average node degree | mean number of edges incident to the nodes |
| intersection count | number of intersections (non-dead-end nodes) in the graph |
| average streets per node | mean number of streets (edges in the undirected representation of the graph) that emanate from each node (intersections and dead-ends) |
| counts of streets per node | a dictionary with keys = the number of streets emanating from the node, and values = the number of nodes with this number |
| proportions of streets per node | a dictionary, same as above, but represents a proportion of the total, rather than raw counts |
| total edge length | sum of all edge lengths in the graph, in meters |
| average edge length | mean edge length in the graph, in meters |
| total street length | sum of all edges in the undirected representation of the graph |
| average street length | mean edge length in the undirected representation of the graph, in meters |
| count of street segments | number of edges in the undirected representation of the graph |
| node density | *n* divided by area in square kilometers |
| edge density | total edge length divided by area in square kilometers |
| street density | total street length divided by area in square kilometers |
| average circuity | total edge length divided by the sum of the great circle distances between the nodes incident to each edge |
| self-loop proportion | proportion of edges that have a single incident node (i.e., the edge links nodes *u* and *v*, and *u*=*v*) |
| average neighborhood degree | mean degree of the nodes in the neighborhood of each node |
| average of the average neighborhood degree | mean of all the average neighborhood degrees in the graph |
| average weighted neighborhood degree | mean degree of the nodes in the neighborhood of each node, weighted by edge length |
| average of the average weighted neighborhood degree | mean of all the weighted average neighborhood degrees in the graph |
| degree centrality | the fraction of nodes that each node is connected to |
| average degree centrality | mean of all the degree centralities in the graph |
| clustering coefficient | for each node, the extent to which nodes tend to cluster together |
| weighted clustering coefficient | for each node, the extent to which nodes tend to cluster together, weighted by edge length |
| average weighted clustering coefficient | mean of the weighted clustering coefficients of all the nodes in the graph |
| pagerank | ranking of nodes based on structure of incoming edges (link analysis) |
| maximum pagerank | the highest pagerank value of any node in the graph |
| maximum pagerank node | the node with the maximum pagerank |
| minimum pagerank | the lowest pagerank value of any node in the graph |
| minimum pagerank node | the node with the minimum pagerank |
| node connectivity | the minimum number of nodes that must be removed to disconnect the graph |
| average node connectivity | the expected number of nodes that must be removed to disconnect a randomly selected pair of non-adjacent nodes |
| edge connectivity | the minimum number of edges that must be removed to disconnect the graph |
| eccentricity | for each node, the maximum distance from it to all other nodes, weighted by length |
| diameter | the maximum eccentricity of any node in the graph |
| radius | the minimum eccentricity of any node in the graph |
| center | the set of all nodes whose eccentricity equals the radius |
| periphery | the set of all nodes whose eccentricity equals the diameter |
| closeness centrality | for each node, the reciprocal of the sum of the distance from the node to all other nodes in the graph, weighted by length |
| average closeness centrality | mean of all the closeness centralities of all the nodes in the graph |
| betweenness centrality | for each node, the fraction of all shortest paths that pass through the node |
| average betweenness centrality | mean of all the betweenness centralities of all the nodes in the graph |

Table 5.1 Network metrics and statistics automatically calculated by OSMnx.

One can also calculate and plot shortest-path routes between points, taking one-way streets into account (Figure 5.7). OSMnx can visualize street segments by length to provide a sense of where a city’s longest and shortest blocks are distributed. OSMnx can similarly visualize one-way vs two-way edges to provide a sense of where a city’s one-way streets and divided roads are distributed. One can also quickly visualize the spatial distribution of cul-de-sacs (or intersections of any type) in a city to get a sense of these points of low network connectivity (Figure 5.8). Allan Jacobs (1995) famously compared several cities’ urban forms through figure-ground diagrams of one square mile of each’s street network in his book Great Streets. We can re-create this automatically and computationally with OSMnx (Figure 5.9). These Jacobsesque figure-ground diagrams are created completely with OSMnx and its network plotting function.

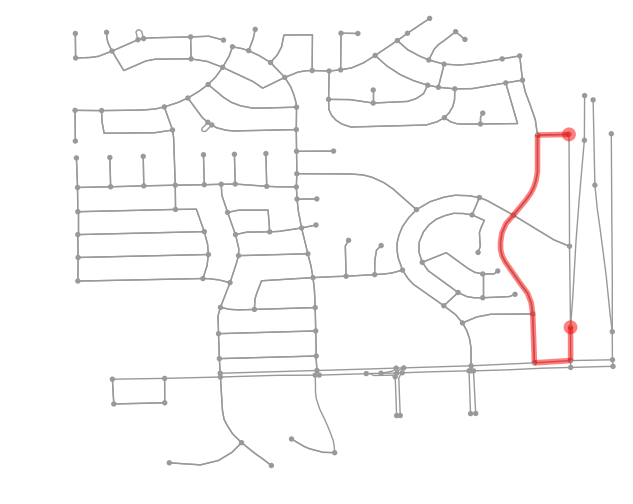


Figure 5.7 OSMnx calculates the shortest network path between two points, accounting for one-way routes, and plots it.



Figure 5.8 OSMnx visualizes the spatial distribution of cul-de-sacs in Piedmont, California.



Figure 5.9 One square mile of each city, created and plotted automatically by OSMnx in the style of Allan Jacobs (1995).

### Summary

To summarize, OSMnx allows researchers and planners to download spatial geometries and construct, project, visualize, and analyze complex street networks. It allows one to automate the collection and computational analysis of street networks for powerful and consistent research, transportation engineering, and urban design. OSMnx is built on top of NetworkX, matplotlib, and geopandas for rich network analytic capabilities, beautiful and simple visualizations, and fast spatial queries with r-tree indexing.

## Discussion

Street network analysis in the urban planning literature suffers from challenges of usability, planarity, reproducibility, and sample sizes. This study created a tool, OSMnx, to make the collection of data and creation and analysis of street networks simple, consistent, and automatable. OSMnx is available on GitHub at: <https://github.com/gboeing/osmnx>.

OSMnx contributes five significant new capabilities for researchers and city planners: first, the automatic downloading of administrative place boundaries and shapefiles; second, the tailored and automated downloading and constructing of street networks from OpenStreetMap; third, the automatic correction and simplification of network topology; fourth, the ability to save street networks to disk as shapefiles, GraphML, or SVG files; and fifth, the ability to analyze street networks, calculate routes, visualize the networks, and calculate network metrics and statistics. These metrics and statistics include both those common in urban design and transportation studies, and metrics that measure the complexity of the network. OSMnx makes the acquisition, construction, and analysis of urban street networks easy, consistent, and reproducible while opening up new data to researchers and practitioners.

# Case Study: Portland, Oregon

## Abstract

This short chapter presents a small case study to simply but plainly demonstrate the usage of OSMnx for research. It collects the street networks for three small subsections of Portland, Oregon to perform a cross-sectional analysis. It introduces these neighborhoods first from a brief qualitative and historical perspective then explores their comparative quantitative measures of network complexity and structure. Then it discusses these empirical findings and what insights may be drawn from them.

## Introduction

As discussed in chapter 5, this study developed a new tool for downloading, constructing, correcting, analyzing, projecting, and visualizing street networks from OpenStreetMap data. OSMnx can work with driving, walking, and biking circulation networks. This chapter presents a small case study to demonstrate the functionality of OSMnx. It examines three small half-kilometer sections of the street network in different neighborhoods in Portland, Oregon. This scale of analysis and sample size are small, but they provide simple, comprehensible examples to illustrate the complex network concepts presented in chapter 3, the network complexity (metric and topological) measures presented in chapter 4, and the methodological tool presented in chapter 5. This chapter serves to tie these threads together empirically and present a visual demonstration before embarking on the large multi-scale analysis of 25,000 street networks in the subsequent chapter, chapter 7.

This chapter has three aims. First, as discussed, it demonstrates the functionality of OSMnx with a simple case study. Second, it presents empirical findings of three street network sections in Portland, Oregon and uses the quantitative measures to compare and contrast these network sections. Third, it offers some insights and suggestions arising from these findings. This chapter is organized as follows. First it lays out the methods by which it uses OSMnx to acquire and analyze these networks. Then it presents the findings from this analysis. Finally, it concludes with a discussion of these findings and their insights.

## Methods

To demonstrate OSMnx, we analyze three neighborhoods in Portland, Oregon. First we define three square bounding boxes of 0.5 km2 each in the city’s downtown, Laurelhurst, and Northwest Heights neighborhoods. The downtown study site is centered on the latitude-longitude coordinates (45.518698, -122.679964), the Laurelhurst study site is centered on the latitude-longitude coordinates (45.527198, -122.625055), and the Northwest Heights study site is centered on the latitude-longitude coordinates (45.539635, -122.770915). These square, half-kilometer study areas are small and do not conform to human definitions of local neighborhood extents, but instead are useful for consistent comparison across sites at a small spatial scale.

These three study sites were selected for their different urban forms and histories. Downtown Portland is geographically constrained between the Willamette River and the West Hills. Its street network, laid out in the late 1800s, is famous for its dense 200-foot regular, orthogonal blocks (Grammenos and Pollard 2009). This fine-grained, walkable network of streets has been heralded by many urban designers as something of an ideal American urban form (e.g., Duany et al. 2010; Speck 2012). Laurelhurst is a primarily residential streetcar suburb in East Portland, developed primarily during the 1910s and 1920s (Snyder 1979). The neighborhood was originally platted by John Charles Olmstead – nephew of Frederick Law Olmstead – and its homes and meandering streets exhibit the character of the American craftsman and garden suburb movements (Guzowski 1990; Works 2016). Northwest Heights is a sprawling neighborhood nestled in the rolling hills west of downtown and north of Beaverton. Developed at the end of the 20th century, it features large, single-family homes and a winding, disconnected street network.

This study uses OSMnx to visualize these street networks then measure them quantitatively to assess their complexity, particularly through the lens of connectedness and robustness. OSMnx downloads the drivable street networks for each square study site, constructs the network, simplifies the topology to make it correct (as discussed in section 5.4.3), and then calculates full network statistics (as described in section 5.4.5).

## Findings

If we project the networks to UTM (zone 10 calculated automatically), and plot them as seen in Figure 5.10, we can get an initial qualitative sense of these small sections of Portland’s street network and how they compare to their disparate histories. OSMnx calculates correct numbers of streets emanating from each intersection (as discussed in section 5.4.5), even for peripheral intersections whose streets were cut off by the bounding box. Also notice in the network of Northwest Heights in Figure 6.1, at the far-right of the panel, there are two nodes that appear to exist in the middle of a street segment, and thus should have been removed during simplification. However, these are actual multi-way intersections that OSMnx properly retained: they simply have a street that would connect to a node outside the right edge of the bounding box.

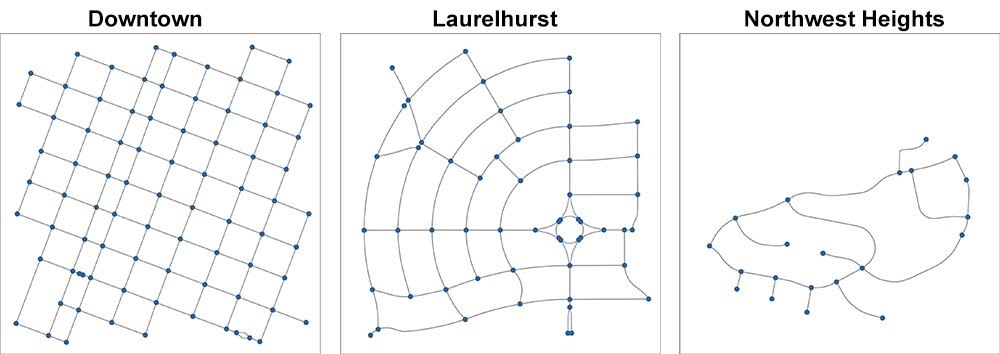


Figure 6.1 Three 0.5 km2 sections (each centered on the coordinates presented in section 6.3) of the street network in Portland, Oregon projected and plotted automatically by OSMnx.

The different histories and designs of these street networks are reflected in Figure 10. These histories and designs represent different historical eras, planning regimes, design paradigms, transportation technologies, and topographies. Several quantitative measures can describe these differences as well (Table 5.2). In terms of density metrics, downtown has 164 intersections/km2, Laurelhurst has 110, and Northwest Heights has 28. Given their respective histories, it is little surprise that downtown has approximately 6x the intersection density of Northwest Heights. Similarly, downtown has 21 linear km of physical street/km2, Laurelhurst has 16, and Northwest Heights has 5. In the downtown network, the total street length equals the total edge length, because every edge is one-way. These values differ somewhat in the other two networks because of the presence of two-way streets. As a linear proxy for block size, the average street segment length is 76 meters in downtown, 92 meters in Laurelhurst, and 117 meters in Northwest Heights. These metric measures tell us quantitatively what we can see by visually inspecting the street networks: downtown’s is fine-grained and dense, Northwest Heights’s is coarse-grained and sparse, and Laurelhurst’s is somewhere in between.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Downtown | Laurelhurst | NW Heights |
| Area (km2) | 0.5 | 0.5 | 0.5 |
| Avg of the avg neighborhood degree | 1.64 | 2.976 | 2.75 |
| Avg of the avg weighted neighborhood degree | 0.024 | 0.059 | 0.03 |
| Avg betweenness centrality | 0.07 | 0.077 | 0.137 |
| Avg circuity | 1.001 | 1.007 | 1.09 |
| Avg closeness centrality | 0.002 | 0.002 | 0.002 |
| Avg clustering coefficient | 0 | 0.108 | 0 |
| Avg weighted clustering coefficient | 0 | 0.023 | 0 |
| Intersection count | 82 | 55 | 14 |
| Avg degree centrality | 0.042 | 0.102 | 0.219 |
| Diameter (m) | 1278.312 | 1021.904 | 898.466 |
| Edge connectivity | 1 | 1 | 1 |
| Edge density (m/km2) | 21315.912 | 29550.502 | 10708.362 |
| Avg edge length (m) | 76.261 | 97.375 | 116.597 |
| Total edge length (m) | 10676.509 | 14801.056 | 5363.461 |
| Intersection density (per km2) | 163.715 | 109.808 | 27.952 |
| Average node degree | 3.415 | 5.527 | 4.381 |
| *m* | 140 | 152 | 46 |
| *n* | 82 | 55 | 21 |
| Node connectivity | 1 | 1 | 1 |
| Avg node connectivity | 1.326 | 2.107 | 1.443 |
| Avg node connectivity (undirected) | 2.868 | 2.496 | 1.443 |
| Node density (per km2) | 163.715 | 109.808 | 41.927 |
| Max pagerank value | 0.03 | 0.029 | 0.106 |
| Min pagerank value | 0.002 | 0.004 | 0.017 |
| Radius (m) | 742.926 | 537.055 | 561.815 |
| Self-loop proportion | 0 | 0 | 0 |
| Street density (m/km2) | 21315.912 | 15575.879 | 5354.181 |
| Average street segment length (m) | 76.261 | 91.783 | 116.597 |
| Total street length (m) | 10676.509 | 7801.541 | 2681.731 |
| Street segment count | 140 | 85 | 23 |
| Average streets per node | 3.927 | 3.582 | 2.381 |

Table 6.1 Descriptive statistics for three street network sections in the city of Portland, Oregon.

Topological measures can tell us more about the complexity, connectivity, and robustness. On average, intersections in the downtown section have 3.9 streets connected to them, in Laurelhurst they have 3.6, and in Northwest Heights they have 2.4 (Table 5.2). Beyond the *average*, the statistical distribution of the number of streets per intersection reveals more about the type of network. Compare Figure 5.10 to Figure 5.11: most intersections downtown are 4-way intersections, whereas Laurelhurst features a nearly even mix of 3-way and 4-way intersections, and Northwest Heights has mostly 3-way intersections and cul-de-sacs (one-third of its nodes are the latter).

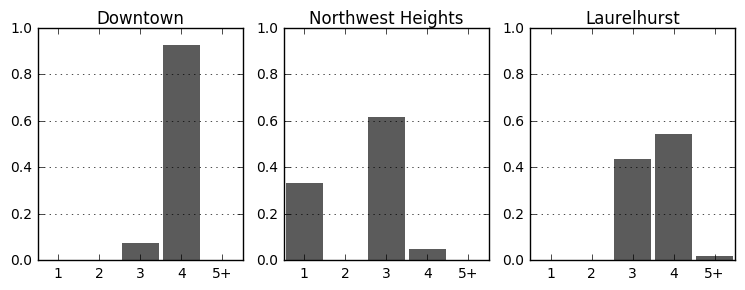


Figure 6.2 Distribution of intersection types in each street network section: *x*-axis is number of streets at the intersection and *y*-axis is the proportion of intersections with that number of streets.

Unsurprisingly (for the reasons discussed in chapter 4), the node and edge connectivity of each network is *1*. More useful is the average node connectivity. Recall that this represents the average number of nodes that must be removed to disconnect a randomly selected pair of non-adjacent nodes. In other words, this is how many non-overlapping paths exist on average between two randomly selected nodes. Thus, on average 1.3 nodes must fail for two nodes to be disconnected in downtown, 2.1 in Laurelhurst, and 1.4 in Northwest Heights (Table 5.2). These values may initially seem surprising: downtown has the *least* robust network despite its density and fine grain. However, this is explained by the fact that *every* street edge in this downtown section is one-way, greatly circumscribing the number of paths between nodes. If we instead examine the *undirected* average node connectivity, it is 2.9 in downtown, 2.5 in Laurelhurst, and 1.4 in Northwest Heights. Thus, were all the edges in all three networks undirected (i.e., two-way streets), downtown’s average node connectivity would more than double and it would have the *most* robust network (by this measure). This finding suggests that there could be considerable complexity, connectivity, and robustness gains in converting downtown’s streets from one-way to two-way.

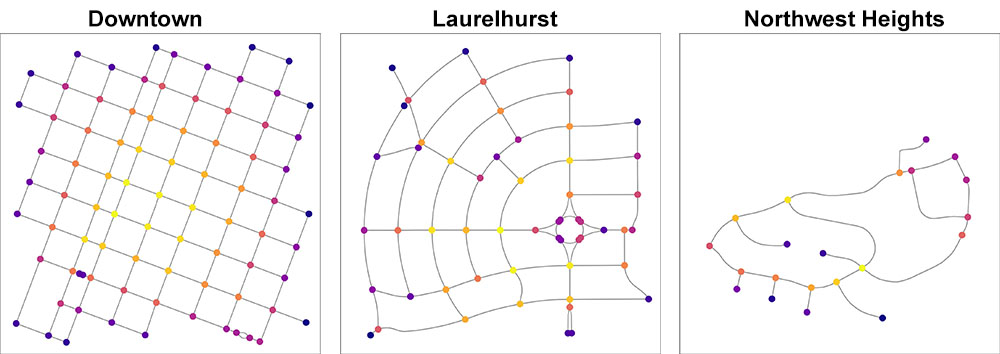


Figure 6.3 Three 0.5 km2 sections of the street network in Portland, Oregon. Nodes colored by betweenness centrality from lowest (dark) to highest (light).

Finally, the average betweenness centrality indicates that 7% of all shortest paths pass through an average node in downtown, 8% in Laurelhurst, and 14% in Northwest Heights (Table 5.2). The spatial distribution of betweenness centralities in these three networks indicates the relative importance of each node (Figure 5.12): the lightest nodes have the most shortest-paths passing through them, and darkest nodes the fewest. In downtown, important nodes are concentrated at the center of the network due to its orthogonality. In Northwest Heights, the two most important nodes are also critical chokepoints connecting the west side of the network to the east. In fact, the most important node in Northwest Heights has 43% of all shortest paths running through it. In contrast, the most important node in downtown has only 15%. The street network in this section of Northwest Heights is thus more prone to disruption if its most important node fails (e.g., due to a traffic jam, flood, or earthquake) than downtown is if its most important node fails.

## Discussion

This chapter presented a short case study demonstrating the abilities of OSMnx to download street network data from OpenStreetMap’s APIs, construct the data into a graph-theoretic network object with NetworkX, analyze the network with various metric and topological measures of its complexity, and project and plot the network to visualize its characteristics and structure.

In the simple case study of Portland, Oregon, we saw how to assess the street network from both metric and topological perspectives using OSMnx. The quantitative analysis corresponded with expectations given the histories of the networks and their visualizations. In particular, we found that these three networks differed substantially in density, connectedness, and robustness. The case study demonstrated how there could be substantial gains in network robustness if one-way streets in the dense, orthogonal downtown were converted to two-way streets. These measures of network structure characterized the complexity of the circulation network in terms of density, permeability, resilience, and connectedness – physical attributes that influence how an urban system structures its interactions, connections, and dynamics.

The next chapter goes beyond this simple demonstration of using the tool to acquire and analyze three small sections of a street network, to instead apply OSMnx empirically in an urban form analysis of 27,000 street networks across the United States at multiple scales.

# Multi-scale analysis of urban street networks

## Abstract

## Introduction

Network analysis is a natural approach to the study of cities as growing, complex systems that self-organize to fill space with a capillary circulation network (Masucci et al. 2009). As discussed in chapter 5, the empirical literature on street network analysis is growing ever richer, but suffers from some limitations. First, sample sizes tend to be fairly small due to data availability, gathering, and processing constraints. Second, reproducibility is difficult when the dozens of decisions that go into analysis – such as spatial extents, topological simplification and correction, definitions of nodes and edges, etc. – are ad hoc or only partly reported. Third, and related to the first two, studies frequently oversimplify to planar or undirected primal graphs for tractability, or use dual graphs despite the loss of geographic and metric information. Fourth, the current landscape of tools and methods offers no ideal technique that balances usability, customizability, reproducibility, and scalability in acquiring, constructing, and analyzing network data.

The previous two chapters in this dissertation specifically addressed the fourth limitation above by introducing OSMnx and demonstrating its use in a small case study of Portland, Oregon. This chapter aims to address the first three limitations by conducting an analysis of street networks at multiple scales, with large sample sizes, with clearly defined network definitions and extents for reproducibility, and using non-planar, directed graphs. In particular, it examines urban street networks – represented here as primal, non-planar, weighted multidigraphs with possible self-loops – through the framework of complexity developed in this dissertation, focusing on structure, connectedness, centrality, and robustness.

Most studies in the street network literature that conduct topological and/or metric analysis tend to have sample sizes ranging around 5 to 50 networks. This chapter instead conducts a large analysis of over 27,000 urban street networks at multiple overlapping scales across the United States. Namely, it examines the street networks of every U.S. incorporated city and town, urbanized area, and Zillow-defined neighborhood. In total, the study presented in this chapter uses OSMnx to download, construct, and analyze 497 urbanized areas’ street networks, 19,655 cities’ and towns’ street networks, and 6,857 neighborhoods’ street networks. It uses these street networks to conduct four analyses: at the metropolitan scale, at the municipal scale, at the neighborhood scale, and a case study looking deeper at the neighborhood-scale street networks in the city of San Francisco.

This chapter is organized as follows. In the next section it presents the data sources, tools, and methods used to collect, construct, and analyze these street networks. Then it presents findings of the analyses at three spatial scales: metropolitan, municipal, and neighborhood. Next it presents a case study of the neighborhood scale in San Francisco. Finally, it concludes with a discussion of these findings and their implications for street network analysis, urban form complexity, and city planning.

## Methods

This study uses OSMnx to download, construct, correct, analyze, and visualize street networks at metropolitan, municipal, and neighborhood scales. To define the study sites and their spatial boundaries, this study uses three sets of geometries. The first defines the metropolitan-scale study sites using the 2016 national TIGER/Line shapefile of US census bureau urban areas, released 19 August 2016. Census-defined urban areas comprise a set of census tracts that meet a minimum density threshold (US Census Bureau 2010). In particular, we retain only the *urbanized areas* subset in this data set (i.e., areas with greater than 50,000 population), discarding the smaller *urban clusters* subset. The second set of geometries defines the municipal-scale study sites using 51 TIGER/Line shapefiles (again, 2016) of US census bureau places for all 50 states plus the District of Columbia. In particular, we discard the subset of census-designated places in this data set, instead retaining all the incorporated cities and towns in the United States. The third and final set of geometries defines the neighborhood-scale study sites using 42 Zillow neighborhood shapefiles. These shapefiles contain the geometries of neighborhoods in 41 states plus the District of Columbia. This is a fairly new data set comprising nearly 7,000 neighborhoods in large American cities, but as Schernthanner et al. (2016) point out, Zillow does not publish the methodology it uses to construct these geometries. However, despite its newness it already has a track record in the academic literature. For instance, Besbris et al. (2015) used Zillow neighborhood boundaries to examine neighborhood stigma, and Albrecht and Abramowitz (2014) used these data to study neighborhood-level poverty in New York.

For all of these geometries, we use OSMnx’s graph\_from\_polygon function to download the street network (drivable plus service roads) contained within the geometry. First it buffers each geometry by 0.5 km, then downloads the street network within this buffered geometry. Next it constructs a street network from this data, corrects the topology, calculates street counts per intersection (this ensures that intersections are not considered cul-de-sacs simply because an incident edge connects to a node outside the desired polygon), then truncates the network to the original, desired polygon – as discussed in chapter 5. OSMnx saves each of these networks to disk as GraphML and shapefiles. Finally, it automatically calculates metric and topological measures for each. For the sake of avoiding unnecessary repetition, these measures will not be defined here, but they were presented in detail in chapters 4 and 5.

In total, this study cross-sectionally analyzes 27,009 networks: 497 urbanized areas’ street networks, 19,655 cities’ and towns’ street networks, and 6,857 neighborhoods’ street networks. These sample sizes are larger than those in any previous study. In the following sections, we present the findings of these analyses at the metropolitan scale, the municipal scale, the neighborhood scale, and through a case study looking deeper at the neighborhood-scale street networks in the city of San Francisco. We examine topological and metric properties as discussed in chapter 4, but in a planar graph, topological and metric properties are somewhat interrelated (Masucci et al. 2009).

## Analysis of Metropolitan-Scale Street Networks

There is great variation in street network characteristics across the entire data set of 497 urbanized areas (Table 7.1). This is unsurprising: urbanized areas span a wide spectrum of sizes, from the Delano, CA Urbanized Area’s 26 km2 to the New York--Newark, NY--NJ--CT Urbanized Area’s 8,937 km2 – thus, density and count-based measures demonstrate substantial variance. Further, these urbanized areas span a wide spectrum of terrains, development eras and paradigms, and cultures.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | mean | std | min | 50% | max |
| area\_km | 460.657 | 858.125 | 25.685 | 184.898 | 8937.429 |
| avg\_neighbor\_degree\_avg | 2.886 | 0.109 | 2.626 | 2.875 | 3.228 |
| avg\_weighted\_neighbor\_degree\_avg | 0.032 | 0.018 | 0.021 | 0.03 | 0.321 |
| circuity\_avg | 1.076 | 0.019 | 1.023 | 1.074 | 1.14 |
| clustering\_coefficient\_avg | 0.042 | 0.009 | 0.015 | 0.042 | 0.071 |
| clustering\_coefficient\_weighted\_avg | 0.002 | 0.001 | 0 | 0.001 | 0.006 |
| count\_intersections | 12582.44 | 26054.18 | 751 | 4593 | 307848 |
| degree\_centrality\_avg | 0.001 | 0.001 | 0 | 0.001 | 0.007 |
| edge\_density\_km | 13455.45 | 2137.624 | 7960.945 | 13351.71 | 21233.61 |
| edge\_length\_avg | 158.588 | 17.653 | 117.341 | 157.332 | 223.08 |
| edge\_length\_total | 6352530 | 12624640 | 427006 | 2392866 | 1.42E+08 |
| geoid | 49207.72 | 28538.9 | 199 | 48799 | 98182 |
| int\_1\_streets\_prop | 0.213 | 0.055 | 0.077 | 0.207 | 0.416 |
| int\_3\_streets\_prop | 0.593 | 0.046 | 0.444 | 0.591 | 0.778 |
| int\_4\_streets\_prop | 0.187 | 0.063 | 0.054 | 0.178 | 0.422 |
| intersection\_density\_km | 26.469 | 6.256 | 12.469 | 26.029 | 49.423 |
| k\_avg | 5.153 | 0.302 | 4.307 | 5.143 | 6.056 |
| m | 40890.09 | 83677.67 | 2516 | 14955 | 981646 |
| n | 16031.73 | 32584.98 | 874 | 5830 | 373309 |
| node\_density\_km | 33.628 | 7.641 | 17.675 | 33.071 | 61.655 |
| pagerank\_max | 0.001 | 0.001 | 0 | 0.001 | 0.003 |
| pagerank\_min | 0 | 0 | 0 | 0 | 0 |
| self\_loop\_proportion | 0.008 | 0.008 | 0 | 0.006 | 0.071 |
| street\_density\_km | 7262.478 | 1220.966 | 4217.416 | 7170.916 | 11796.75 |
| street\_length\_avg | 161.331 | 17.765 | 119.573 | 160.288 | 225.92 |
| street\_length\_total | 3479997 | 7026153 | 222327.9 | 1269337 | 79046836 |
| street\_segments\_count | 22011.37 | 45725.34 | 1281 | 7868 | 533757 |
| streets\_per\_node\_avg | 2.764 | 0.162 | 2.223 | 2.770 | 3.217 |

Table 7.1 Measures of central tendency and dispersion for selected measures of the 497 urbanized area street networks.

Nevertheless, looking across the data set provides a sense of the breadth of American metropolitan street networks. New York’s urbanized area – America’s largest – has 373,309 intersections and 79 million meters of linear street (or 417,570 and 83.4 million if you include service roads). Delano, CA’s urbanized area – America’s smallest – has 874 intersections and 222,328 meters of linear street (or 964 and 231,000 meters if you include service roads). The typical (Table 7.1, median) American urbanized area is approximate 185 km2 (land area), has 5,830 intersections, and 1.3 million meters of street. Its street network is about 7.4% more circuitous than straight-line edges between nodes would be. The most circuitous network is 14% more circuitous than straight-line would be, and least is only 2%. Looking at network complexity in terms of density and connectivity, in the typical urbanized area, the average street segment length (a proxy for block size) is 160 meters. The longest average street segment is the 226 m average of urbanized Danbury, CT. Puerto Rican cities hold the top four positions for shortest average street segment length, but among the 50 states plus DC, the shortest average street segment is the 125.3 m average of urbanized Tracy, CA, indicating a much finer street network. The urbanized area of Portland, Oregon, with its famously compact walkable blocks, ranks 2nd at 125.5 m on average.

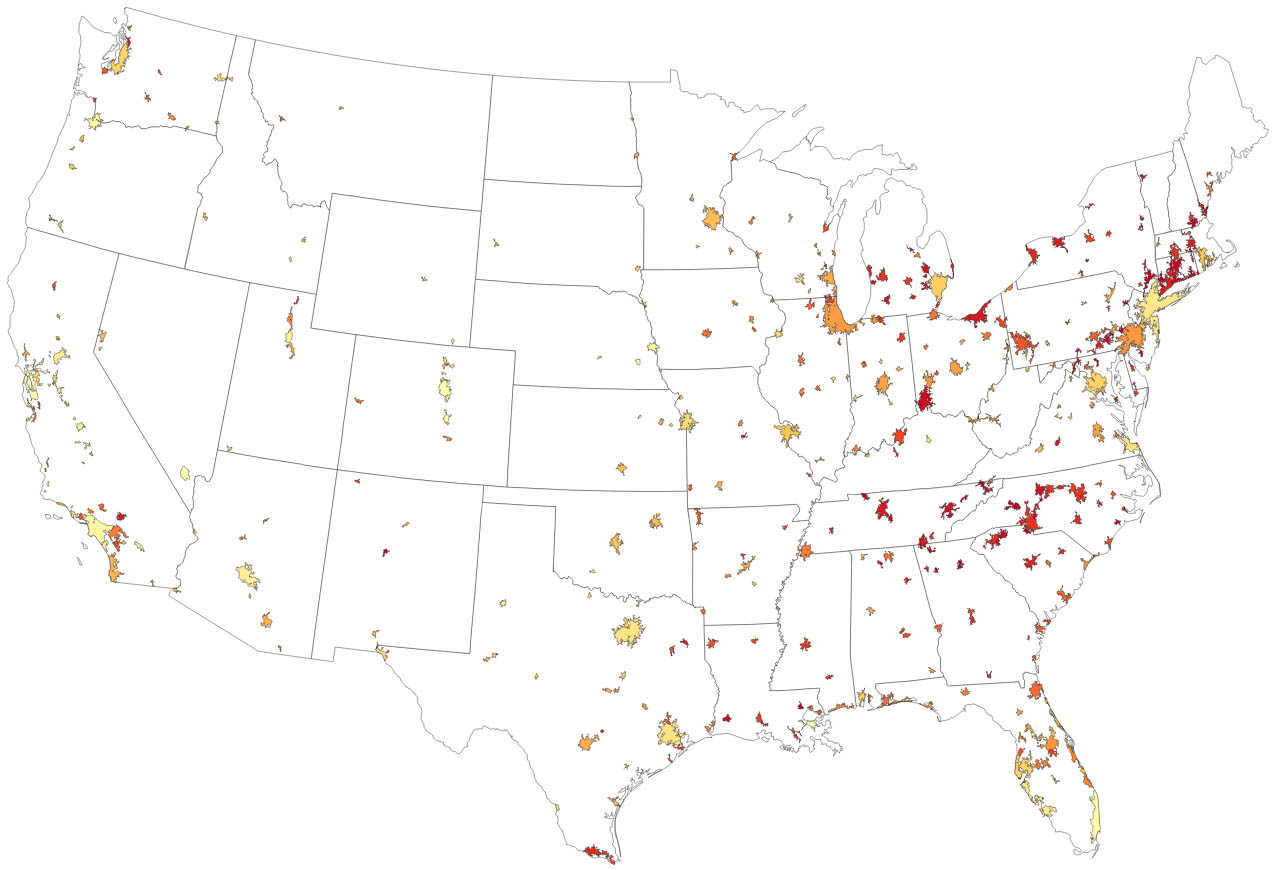


Figure 7.1 Intersection density per urbanized area, from lowest (dark red) to highest (light yellow), in the contiguous US.

The typical urbanized area has 26 intersections per km2. Both the densest and the sparsest are in the deep south: the sparsest is 12.5 (Gainesville, GA urbanized area) and the densest is 49.4 (New Orleans urbanized area). However, New Orleans is an anomaly in the deep south. Figure 7.1 depicts the intersection density in each American urbanized area, from lowest density in dark red to highest density in light yellow. The map makes clear how the highest intersection densities are concentrated to the west of the Mississippi River. The lowest intersection densities are concentrated in a belt running from Louisiana, up through the Carolinas and Appalachians, and into New England. Only the largest cities on the east coast (e.g., Boston, New York, Philadelphia, Washington) and coastal Florida escape this trend.

The distribution of node types (i.e., intersections and dead ends) provides a clear indicator of network connectedness. The typical urbanized area has 2.8 streets per intersection on average: lots of 3-way intersections, fewer cul-de-sacs, and even fewer 4-way intersections. The grid-like San Angelo, TX urbanized area has the most streets connected to each intersection (3.2) on average, and (outside of Puerto Rico, which contains the seven lowest urbanized areas) the sprawling, disconnected Lexington Park, MD urbanized area has the fewest (2.2). These two urban areas fit the trend seen in the spatial distribution across the US: urbanized areas in the great plains and upper Midwest have particularly high numbers of streets per intersection on average, indicating more grid-like, connected networks. Cities in the southern and western US tend to have fewer streets per intersection, reflecting more dead-ends and a disconnected network. This finding is discussed in more detail in the upcoming section.

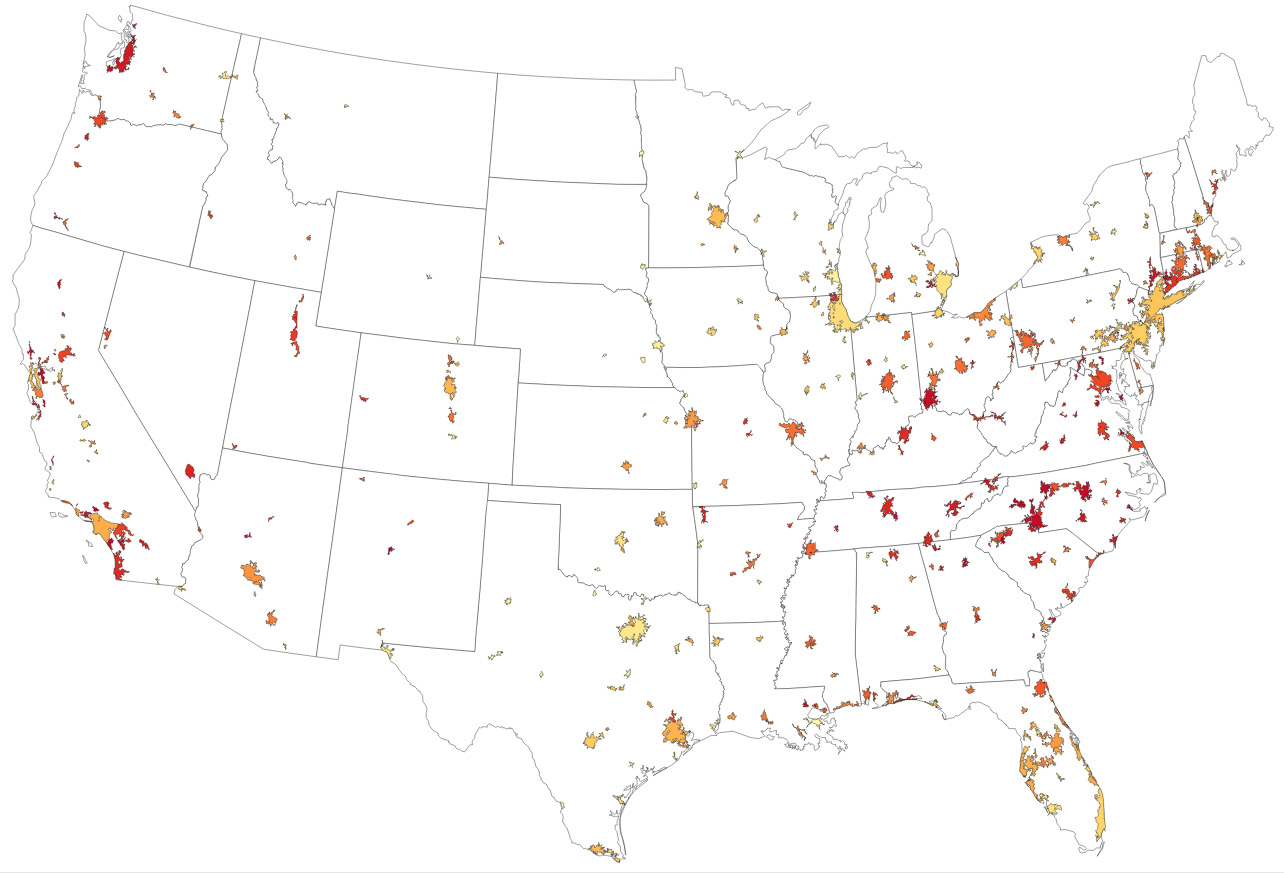


Figure 7.2 Average streets per intersection per urbanized area, from fewest (dark red) to most (light yellow), in the contiguous US.



Figure 7.3 Distribution of node types in 9 urbanized areas, with number of streets emanating from the node on the *x*-axis and proportion of nodes of this type on the *y*-axis (cf. Figure 6.2).

In the typical urbanized area, 18% of nodes are 4-way intersections, 59% are 3-way intersections, and 21% are dead-ends. However, this distribution varies substantially between urbanized areas. Examining a small sample of 9 urbanized areas, chosen to maximize variance, reveals this in clearer detail. In Figure 7.2, urban Atlanta and Chattanooga have very high proportions of dead-ends, each over 30% of all nodes, and very few 4-way intersections, indicating a disconnected street pattern. The urban areas of Phoenix, Boston, Detroit, and Chattanooga have particularly high proportions of 3-way intersections, each over 60%, indicating a prevalence of T-intersections. Conversely, Chicago, New Orleans, Duluth, and Lubbock have high proportions of 4-way intersections, indicating more grid-like connected networks. The relationship between fine-grained networks and connectedness/grid-ness is, however, not clear-cut: intersection density has only a weak, positive linear relationship with the proportion of 4-way intersections in the urbanized area (r2=0.17), as seen in Figure 7.3. But the relationship between network circuity and grid-ness is somewhat clearer: average circuity has a negative linear relationship with the proportion of 4-way intersections in the urbanized area (r2=0.43).

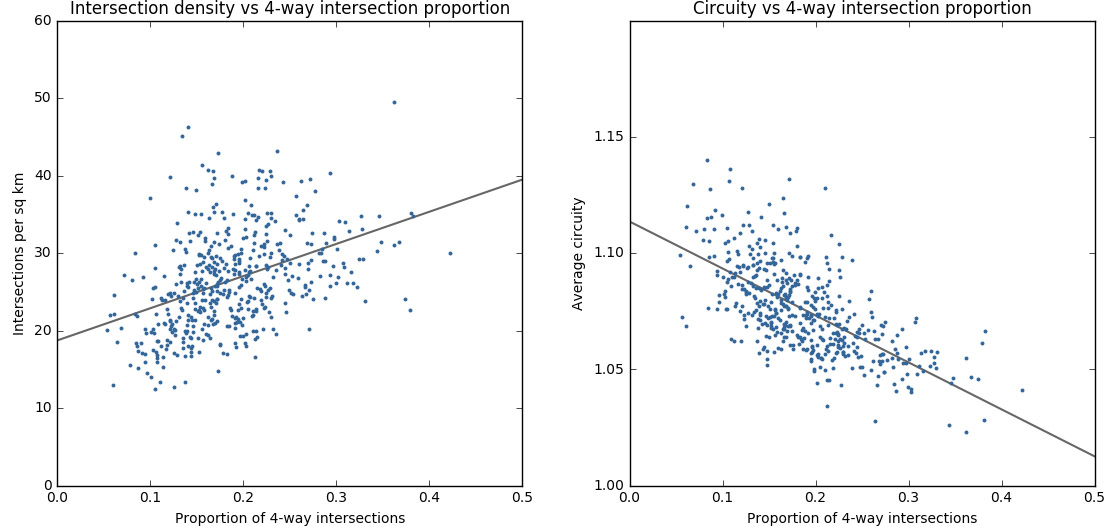


Figure 7.4 Scatterplots of intersection density vs 4-way intersection proportion (left) and average circuity vs 4-way intersection proportion (right), with simple regression lines to indicate the trend.

Focusing on the 30 largest urbanized areas, we can look cross-sectionally at how their metric and topological measures compare (Table 7.2). Among these urban areas, Milwaukee has the least circuitous network (6% more circuitous than straight-line edges would be), and Orlando has the most (12%). San Juan and Atlanta have the fewest streets per node on average (2.36 and 2.45, respectively), while Milwaukee has the most (3.03). Cincinnati has both the lowest intersection density (18 per km2) and street density (6.1 km per km2) while Denver has the highest intersection density (40.6 per km2) and Miami and Los Angeles have the highest street density (10.6 km per km2 apiece). In other words, Cincinnati has a particularly coarse-grained network with few connections and paths. This can also be seen in the average street segment length (a proxy for block size): Cincinnati has the 2nd highest (186 m), bested only by Cleveland (198 m). In contrast, the two lowest are Denver’s 138 m average and San Juan’s 131 m average.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Core  city | Land area km2 | Avg  circuity | Avg cluster  coeff | Cul-de-sac  ratio | 3-way  ratio | 4-way  ratio | Intersect  density (per km2) | Street  density (m/km2) | Avg street  length (m) | Avg streets  per intersect |
| New York | 8937.43 | 1.06 | 0.04 | 0.18 | 0.62 | 0.2 | 34.44 | 8844.47 | 148.1 | 2.86 |
| Atlanta | 6850.05 | 1.1 | 0.04 | 0.32 | 0.58 | 0.09 | 18.39 | 6163.99 | 185.66 | 2.45 |
| Chicago | 6325.26 | 1.07 | 0.04 | 0.17 | 0.57 | 0.25 | 27.05 | 7772.81 | 163.25 | 2.92 |
| Philadelphia | 5132.1 | 1.08 | 0.05 | 0.17 | 0.63 | 0.2 | 26.65 | 7302.51 | 159.19 | 2.87 |
| Boston | 4852.29 | 1.09 | 0.05 | 0.2 | 0.68 | 0.11 | 24.23 | 6435.44 | 153.96 | 2.71 |
| Dallas | 4612.32 | 1.07 | 0.05 | 0.15 | 0.61 | 0.23 | 34.16 | 9157.1 | 155.61 | 2.95 |
| Los Angeles | 4496.6 | 1.06 | 0.03 | 0.21 | 0.56 | 0.22 | 39.45 | 10589.61 | 150.87 | 2.82 |
| Houston | 4303.34 | 1.08 | 0.04 | 0.2 | 0.57 | 0.22 | 33.49 | 8622.66 | 145.02 | 2.83 |
| Detroit | 3461.21 | 1.07 | 0.04 | 0.15 | 0.63 | 0.22 | 31.1 | 8560.97 | 159.24 | 2.93 |
| Washington | 3424.25 | 1.09 | 0.04 | 0.26 | 0.56 | 0.17 | 31.22 | 8255.2 | 145.77 | 2.66 |
| Miami | 3203.52 | 1.1 | 0.05 | 0.17 | 0.59 | 0.23 | 40.54 | 10619.55 | 148.88 | 2.89 |
| Phoenix | 2968.28 | 1.09 | 0.05 | 0.2 | 0.62 | 0.17 | 35.31 | 9099.78 | 149.5 | 2.77 |
| Minneapolis | 2646.6 | 1.08 | 0.05 | 0.19 | 0.57 | 0.23 | 29.54 | 8675.71 | 166.55 | 2.84 |
| Seattle | 2616.71 | 1.07 | 0.03 | 0.3 | 0.54 | 0.16 | 31.57 | 8198.39 | 141.91 | 2.57 |
| Tampa | 2478.76 | 1.1 | 0.05 | 0.2 | 0.58 | 0.21 | 31.35 | 8457.73 | 153.1 | 2.83 |
| St. Louis | 2392.11 | 1.1 | 0.04 | 0.22 | 0.62 | 0.15 | 29.68 | 8160.22 | 154.41 | 2.73 |
| Pittsburgh | 2344.64 | 1.09 | 0.04 | 0.23 | 0.6 | 0.16 | 23.57 | 6709.04 | 164.64 | 2.72 |
| San Juan | 2244.93 | 1.11 | 0.02 | 0.36 | 0.56 | 0.08 | 26.57 | 6427.24 | 131.44 | 2.36 |
| Cincinnati | 2040.25 | 1.07 | 0.03 | 0.31 | 0.54 | 0.14 | 17.96 | 6099.45 | 185.76 | 2.51 |
| Cleveland | 2003.99 | 1.07 | 0.04 | 0.19 | 0.66 | 0.14 | 19.13 | 6508.46 | 197.58 | 2.76 |
| Charlotte | 1919.69 | 1.08 | 0.04 | 0.3 | 0.57 | 0.11 | 21 | 6431.19 | 170.13 | 2.51 |
| San Diego | 1897.37 | 1.08 | 0.03 | 0.28 | 0.54 | 0.17 | 28.89 | 8320.49 | 159.4 | 2.62 |
| Baltimore | 1856.86 | 1.09 | 0.04 | 0.23 | 0.59 | 0.17 | 27.72 | 7559.63 | 152.32 | 2.72 |
| Indianapolis | 1827.96 | 1.08 | 0.05 | 0.23 | 0.59 | 0.17 | 27.62 | 7630.07 | 156.98 | 2.7 |
| Kansas City | 1755.83 | 1.06 | 0.04 | 0.21 | 0.58 | 0.2 | 32.09 | 8568.44 | 151.71 | 2.79 |
| Denver | 1729.19 | 1.07 | 0.05 | 0.2 | 0.57 | 0.22 | 40.6 | 9844.94 | 138.05 | 2.84 |
| Orlando | 1547.96 | 1.11 | 0.06 | 0.2 | 0.61 | 0.18 | 26.3 | 7441.59 | 163.34 | 2.79 |
| San Antonio | 1546.75 | 1.07 | 0.05 | 0.17 | 0.6 | 0.21 | 28.33 | 7910.36 | 162.45 | 2.87 |
| Nashville | 1459.61 | 1.08 | 0.03 | 0.27 | 0.59 | 0.14 | 19.08 | 6102.07 | 181.1 | 2.6 |
| Milwaukee | 1413.45 | 1.06 | 0.06 | 0.14 | 0.55 | 0.3 | 28.27 | 7809 | 157.49 | 3.03 |

Table 7.2 Selected street network stats for the 30 largest urbanized areas.

## Analysis of Municipal-Scale Street Networks

There is similarly great variation in street network characteristics across the entire data set of 19,655 cities and towns (Table 7.3). Again, this data set comprises the street networks of every incorporated city and town in the United States. Following the recent work by Barthelemy and Flammini (2008) and Strano et al. (2013), we examine the relationship between the total street length *L* and the number of nodes *n*. Barthelemy and Flammini proposed a model of cities in which *L* and *n* scale as *n*1/2 and Strano et al. claimed to confirm this finding empirically with a small sample of ten European cities’ street networks. However, their extremely small sample size gives us pause. To investigate this, we examine the relationship between the total street length *L* and the number of nodes *n* for 19,655 US cities and towns and find a clear linear relationship (r2=0.98), contradicting Strano et al., as depicted in Figure 7.5.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | mean | std | min | 50% | max |
| area\_km | 16.703 | 107.499 | 0.039 | 3.918 | 7434.258 |
| avg\_neighbor\_degree\_avg | 2.94 | 0.297 | 0.4 | 2.953 | 3.735 |
| avg\_weighted\_neighbor\_degree\_avg | 0.033 | 0.141 | 0 | 0.029 | 9.357 |
| circuity\_avg | 1.067 | 0.159 | 1 | 1.055 | 20.452 |
| clustering\_coefficient\_avg | 0.048 | 0.041 | 0 | 0.04 | 1 |
| clustering\_coefficient\_weighted\_avg | 0.01 | 0.018 | 0 | 0.005 | 0.524 |
| count\_intersections | 323.834 | 1266.253 | 0 | 83 | 62996 |
| degree\_centrality\_avg | 0.093 | 0.136 | 0 | 0.052 | 2.667 |
| edge\_density\_km | 12654.97 | 6705.182 | 5.602 | 11814.33 | 58603.42 |
| edge\_length\_avg | 161.184 | 80.769 | 25.822 | 144.447 | 3036.957 |
| edge\_length\_total | 159067 | 578521.2 | 51.644 | 40985.86 | 24728326 |
| int\_1\_streets\_prop | 0.192 | 0.093 | 0 | 0.184 | 1 |
| int\_3\_streets\_prop | 0.572 | 0.11 | 0 | 0.579 | 1 |
| int\_4\_streets\_prop | 0.237 | 0.129 | 0 | 0.217 | 1 |
| intersection\_density\_km | 29.363 | 21.607 | 0 | 24.719 | 259.647 |
| k\_avg | 5.251 | 0.668 | 0.8 | 5.268 | 7.166 |
| m | 1046.358 | 3923.863 | 2 | 275 | 176161 |
| n | 400.666 | 1515.772 | 2 | 103 | 71993 |
| node\_density\_km | 35.449 | 24.409 | 0.047 | 30.718 | 296.74 |
| pagerank\_max | 0.034 | 0.046 | 0 | 0.021 | 0.87 |
| pagerank\_min | 0.005 | 0.018 | 0 | 0.002 | 0.5 |
| self\_loop\_proportion | 0.005 | 0.015 | 0 | 0 | 1 |
| street\_density\_km | 6527.859 | 3434.689 | 2.801 | 6109.256 | 29301.71 |
| street\_length\_avg | 162.408 | 81.035 | 25.822 | 145.479 | 3036.957 |
| street\_length\_total | 86096.21 | 331048.8 | 25.822 | 21004.79 | 15348008 |
| street\_segments\_count | 557.52 | 2208.18 | 1 | 140 | 107393 |
| streets\_per\_node\_avg | 2.851 | 0.282 | 1 | 2.852 | 4 |

Table 7.3 Selected summary stats for every city and town in the US

Moreover, previous findings (e.g., Masucci et al. 2009) suggest the distribution of street segment lengths in an urban street network follows a power-law distribution. However, we find that these networks by and large follow a lognormal distribution instead (see theoretical discussion in chapter 3). This finding makes sense: street networks are not scale-free. While there are few very long street segments (say, 1000 m), more medium-length segments (say, 250 m), and lots of short segments (say, 80 m), there are very few very short segments (say, 10 m). This theoretical depiction suggests the lognormal distribution typically found across these US street networks. An exception, of course, lies in consistent, orthogonal grids filling a city’s municipal limits. Such distributions are extremely peaked around a single value – the typical linear length of a grid block.

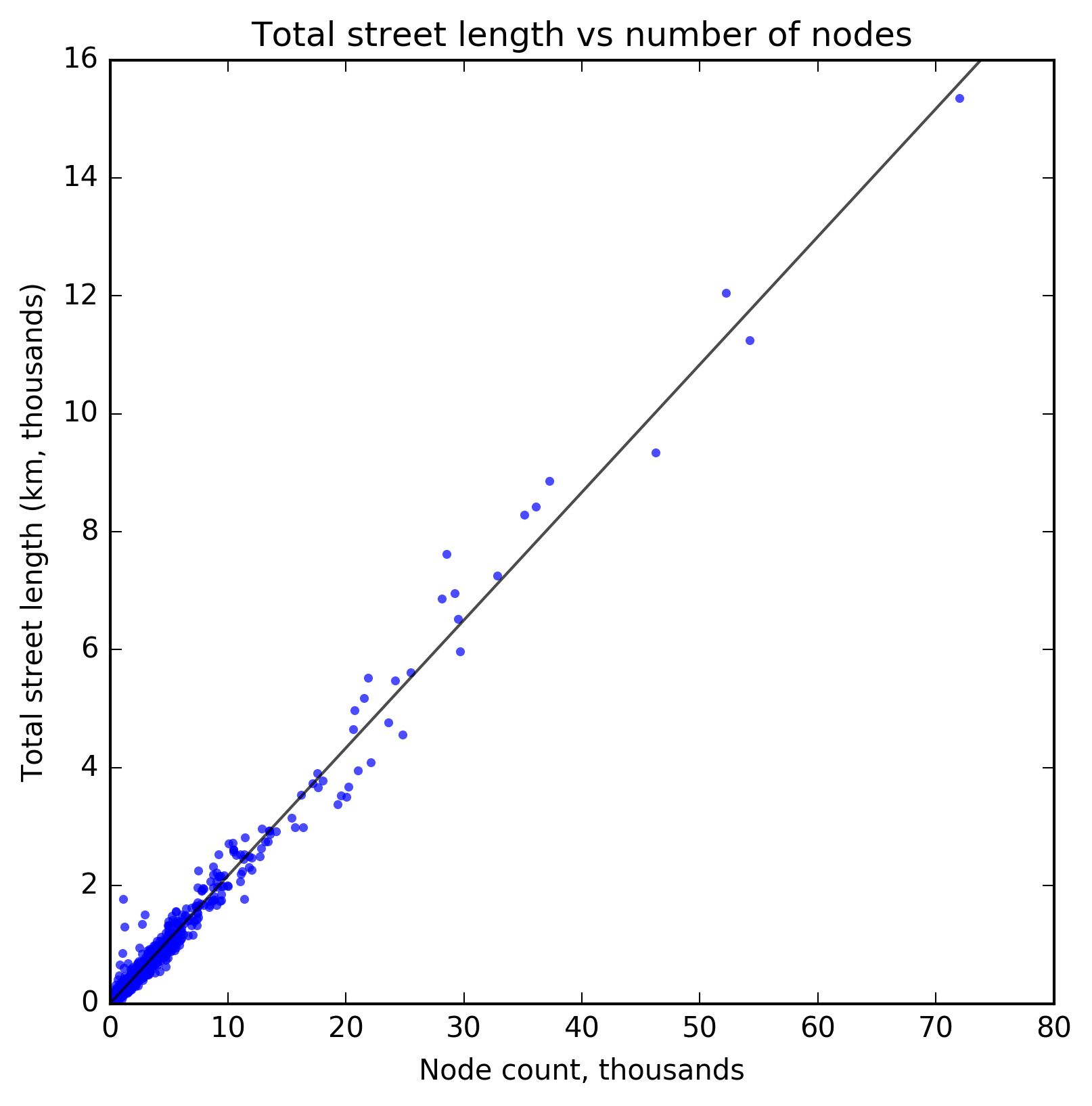


Figure 7.5 The linear relationship between total street length and number of nodes for 19,655 US cities and towns.

Such cities are not uncommon, particularly between the Mississippi River and the Rocky Mountains. These Great Plains states are characterized by a unique street network form that is both orthogonal and reasonably dense. The former is partly the result of topography (flat, plains terrain that allows idealized grids) and design history (platting and development during the late 19th century) that favor orthogonal grids. The latter seems to result from the fact that most towns across the Great Plains exhibit minimal suburban sprawl. Thus, the municipal boundaries snugly embrace the grid-like street network, without extending to accommodate 20th century sprawl, circuity, and loops and lollipops.

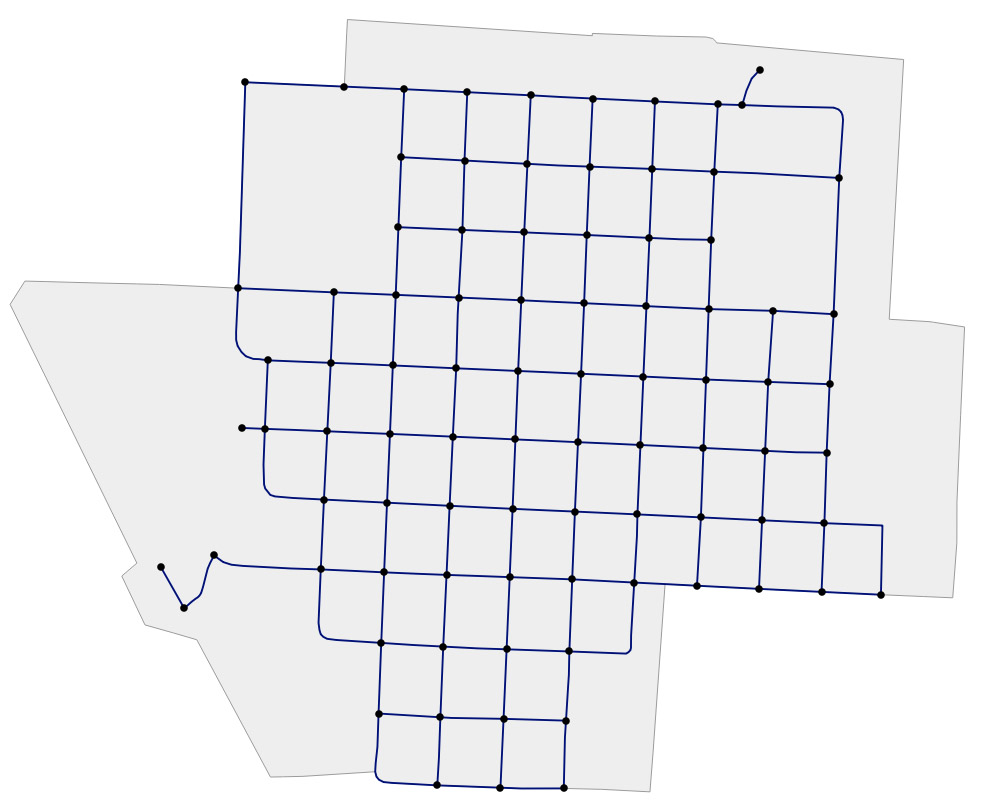


Figure 7.6 Orleans, Nebraska exhibits a compact grid-like street network typical of the Great Plains towns. Municipal extents shown in gray.

For example, if we measure connectedness in terms of *average number of streets per node* at the city-scale and then aggregate these cities by state (Table 7.4), we find Nebraska, Kansas, South Dakota, Montana, North Dakota, Oklahoma, and Iowa have, in order, the highest medians (Figure 7.7). This indicates the most grid-like networks. If we measure density and connectedness in terms of *intersection density* – previously identified as an emergent property of complex network organization – at the city-scale and then aggregate these cities by state, we find Rhode Island, Nebraska, New Jersey, Kansas, and Montana have, in order, the highest medians. We again see three Great Plains states near the top, alongside densely populated East Coast states. Nebraska also has the smallest block sizes (measured with the proxy of average street segment length) while the largest are concentrated in the deep South, upper New England, and Utah.

Municipal boundaries vary greatly in their extents around the built-up area. For example, while Rhode Island averages 56 intersections/km2 in its cities and towns, Alaska averages only 1.3. This is an artifact of Alaska’s municipal boundaries often extending thousands of km2 beyond the built-up area. In fact, Alaska has four *cities* (Anchorage, Juneau, Sitka, and Wrangell) with such large municipal extents that their land areas exceed that of the *state* of Rhode Island.

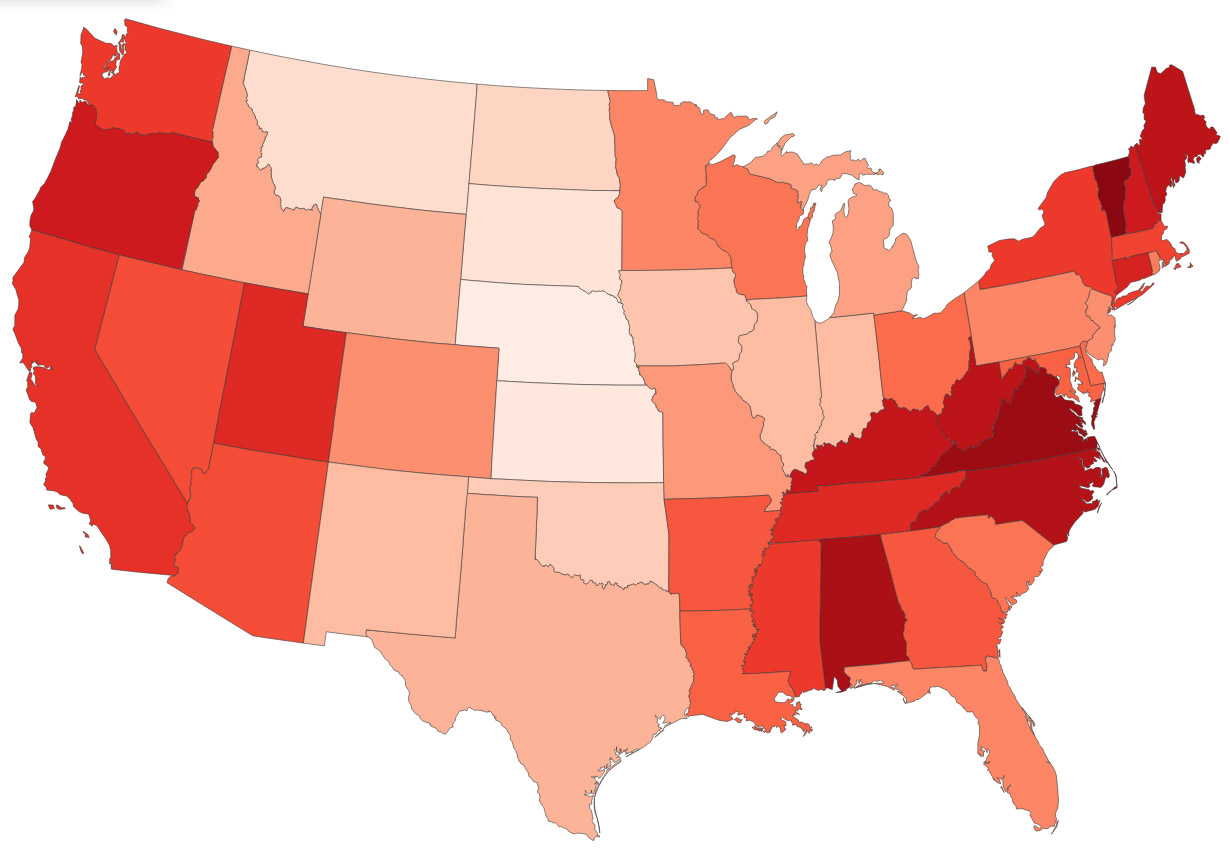


Figure 7.7 Contiguous US states by median average number of streets per node in city and town street networks, colored from lowest/least-connected (dark red) to highest/most-connected (light orange).

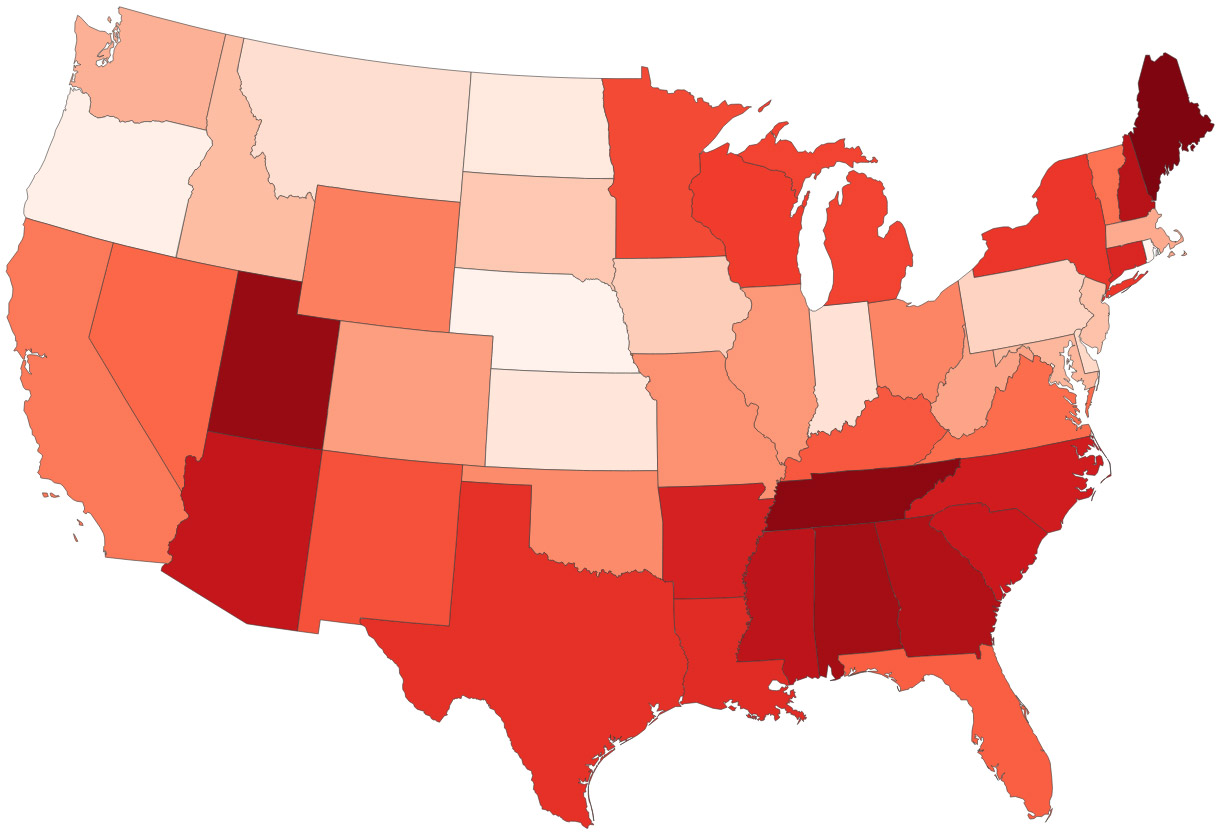


Figure 7.8 Contiguous US states by median average street segment length in city and town street networks, colored from longest/coarsest-grain (dark red) to shortest/finest-grain (light orange).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| state | intersection\_density\_km | streets\_per\_node\_avg | circuity\_avg | street\_length\_avg |
| AK | 1.28 | 2.43 | 1.1 | 223.5 |
| AL | 9.7 | 2.64 | 1.07 | 190.81 |
| AR | 15.75 | 2.78 | 1.06 | 166.32 |
| AZ | 12.45 | 2.77 | 1.08 | 171.8 |
| CA | 32.58 | 2.74 | 1.07 | 143.79 |
| CO | 29.26 | 2.88 | 1.06 | 136.68 |
| CT | 28.05 | 2.7 | 1.07 | 165.87 |
| DC | 58.91 | 3.26 | 1.04 | 122.23 |
| DE | 25.3 | 2.8 | 1.06 | 127.8 |
| FL | 26.26 | 2.87 | 1.07 | 150.75 |
| GA | 15.25 | 2.78 | 1.07 | 177.5 |
| HI | 8 | 2.42 | 1.07 | 177.93 |
| IA | 24.08 | 3.02 | 1.04 | 129.36 |
| ID | 33.85 | 2.91 | 1.06 | 132.08 |
| IL | 29.02 | 2.93 | 1.05 | 137.77 |
| IN | 35.25 | 2.93 | 1.05 | 125.72 |
| KS | 43.94 | 3.14 | 1.04 | 124.39 |
| KY | 25.12 | 2.68 | 1.07 | 151.28 |
| LA | 17.14 | 2.79 | 1.06 | 162.62 |
| MA | 32.33 | 2.76 | 1.07 | 135.98 |
| MD | 28.67 | 2.79 | 1.07 | 133.69 |
| ME | 7.69 | 2.67 | 1.07 | 198.93 |
| MI | 20.93 | 2.9 | 1.05 | 153.5 |
| MN | 18.96 | 2.87 | 1.06 | 152.92 |
| MO | 29.87 | 2.89 | 1.06 | 138.29 |
| MS | 14.76 | 2.75 | 1.06 | 174.86 |
| MT | 38.94 | 3.11 | 1.04 | 126.89 |
| NC | 19.28 | 2.65 | 1.06 | 166.69 |
| ND | 34.28 | 3.07 | 1.04 | 123.93 |
| NE | 45.89 | 3.16 | 1.04 | 119.79 |
| NH | 12.22 | 2.69 | 1.1 | 175.88 |
| NJ | 44.98 | 2.88 | 1.04 | 130.79 |
| NM | 18.5 | 2.93 | 1.05 | 152.02 |
| NV | 13.86 | 2.77 | 1.07 | 147.35 |
| NY | 21.89 | 2.75 | 1.06 | 156.88 |
| OH | 25.23 | 2.8 | 1.05 | 142.08 |
| OK | 28.22 | 3.03 | 1.05 | 139.5 |
| OR | 35.08 | 2.69 | 1.06 | 121.18 |
| PA | 35.69 | 2.87 | 1.05 | 128.34 |
| RI | 56.23 | 2.86 | 1.05 | 110.35 |
| SC | 18.76 | 2.81 | 1.06 | 169.21 |
| SD | 32.01 | 3.12 | 1.04 | 130.75 |
| TN | 13.62 | 2.71 | 1.07 | 192.83 |
| TX | 23.85 | 2.92 | 1.05 | 160.44 |
| UT | 12.58 | 2.71 | 1.06 | 191.04 |
| VA | 25.18 | 2.63 | 1.08 | 145.65 |
| VT | 18.91 | 2.55 | 1.08 | 145.18 |
| WA | 28.71 | 2.75 | 1.06 | 134.02 |
| WI | 17.87 | 2.81 | 1.06 | 156.19 |
| WV | 28.45 | 2.67 | 1.08 | 136.57 |
| WY | 23.48 | 2.92 | 1.06 | 143.63 |

Table 7.5 Median values, aggregated by state plus DC, of selected measures of the municipal-scale street networks for every city and town in the US.

## Analysis of Neighborhood-Scale Street Networks

## Intrametropolitan Neighborhood-Scale Analysis of San Francisco

In Ocean View, the center has low betweenness centrality due to its disconnectivity. However, in the Mission, the center has high betweenness centrality due to its orthogonal gridlike connectivity.

|  |  |
| --- | --- |
| Neighborhood | Max betweenness centrality |
| Outer Sunset | 0.096 |
| Diamond Heights | 0.113 |
| Mission | 0.146 |
| Seacliff | 0.15 |
| Pacific Heights | 0.15 |
| Bayview | 0.161 |
| Nob Hill | 0.165 |
| Inner Richmond | 0.174 |
| Marina | 0.186 |
| Parkside | 0.193 |
| Lakeshore | 0.194 |
| Bernal Heights | 0.199 |
| Outer Richmond | 0.204 |
| Excelsior | 0.206 |
| Downtown | 0.207 |
| Glen Park | 0.209 |
| Western Addition | 0.213 |
| Financial District | 0.216 |
| Noe Valley | 0.218 |
| West Of Twin Peaks | 0.232 |
| Castro-Upper Market | 0.241 |
| Russian Hill | 0.254 |
| Potrero Hill | 0.261 |
| Visitacion Valley | 0.262 |
| Inner Sunset | 0.271 |
| Outer Mission | 0.279 |
| Ocean View | 0.285 |
| Haight-Ashbury | 0.286 |
| Presidio Heights | 0.292 |
| Crocker Amazon | 0.301 |
| South Of Market | 0.31 |
| North Beach | 0.317 |
| Chinatown | 0.362 |
| Twin Peaks | 0.371 |

Table 7.6 San Francisco neighborhoods, by maximum node betweenness centrality

## Discussion

A critical takeaway of this analysis is that scale matters. The characteristics of an urban network for a city fundamentally depend on what *city* means: municipal boundaries, urbanized areas, or just certain neighborhoods? The first is a merely legal definition, but also captures the scope of city planning authority and decision-making for top-down interventions into a street network. The second captures a wider self-organized human system and its emergent built form, but tends to aggregate multiple non-homogenous built forms together into a single unit of analysis. The third captures the nature of the local built form and lived experience, but at the expense of a broad view of the urban system and metropolitan-scale trip-taking. In short, multiple scales in concert provide a clearer view of the urban form and the topological and metric complexity of the street network than any single scale can.

Comparing median street networks of each state, Nebraska has the lowest circuity, the highest average number of streets per node, the second shortest average street segment length, and the second highest intersection density.

# Conclusion

## Synopsis of the Dissertation

Cities – human ecosystems – are complex systems comprising many interrelated, interacting components. Humans both shape their ecosystems (i.e., institutions, cultures, physical built environments), and are in turn shaped by them. Urban complexity is manifested through the self-organization of these components and the emergence of large-scale structure and characteristics. In particular, these components interact through networks – both virtual (Internet and telecommunication, flows of capital, etc.) and physical (street networks, rail networks, etc.). This dissertation focused on street networks and their complexity, emphasizing density, permeability, resilience, and connectedness. These attributes influence the way an urban system’s physical links can structure complex interactions, connections, and dynamics.

This dissertation developed its theoretical framing over chapters 2 and 3. Chapter 2 introduced the background of the nonlinear paradigm, discussing systems, dynamics, self-similarity, and prediction. These foundations set up the theoretical frame in chapter 3 of complexity, cities, and the study of networks. Chapter 4 collated various measures of complexity from the literature into a typology of measures of the complexity of urban form, emphasizing the scale of urban design. In particular, it presented several measures of network complexity that were then operationalized in Chapter 5, 6, and 7. This chapter introduced OSMnx, a new tool to acquire, construct, correct, visualize, and analyze complex urban street networks. Chapter 6 then applied OSMnx empirically in a small case study of street networks in Portland, Oregon as a demonstration of its usage. Chapter 7 then expanded the empirical application of OSMnx to a large study of 27,000 urban street networks at various scales.

## Summary of Key Contributions

### Contribution to the Literature

This dissertation makes various theoretical, methodological, and empirical contributions across its six substantive chapters to the urban planning research literature. This subsection addresses them in order.

Chapter 2 makes one methodological and one theoretical contribution to the planning literature. The former contribution is methodological as this chapter presents Pynamical, a new tool to visualize and explore nonlinear dynamical systems’ behavior. Comparable tools usually must be developed from scratch or rely on expensive commercial software such as MATLAB. Developing tools for exploring, understanding, and visualizing dynamical systems in Python makes them available to a much wider audience of systems analysts, researchers, and students. Pynamical provides a fast, simple, reusable, extensible, free, and open-source new means for exploring system behavior – particularly for the qualitative analysis of such systems in research and pedagogy.

The latter contribution is not original; rather it reviews the theory of nonlinearity for the qualitative analysis of nonlinear dynamical systems’ behavior to an interdisciplinary body of urban scholars and planners. Most formal treatments of chaos and nonlinear dynamics in the scholarly literature are densely technical and geared toward an exclusive audience of mathematicians and physicists. Instead, chapter 2 offers a gentle step-by-step introduction to dynamical systems to a broad social science audience to provide a strong and unambiguous footing for forays into complexity studies. Similarly, chapter 3 offers a minor theoretical contribution to the planning literature by by unpacking the key foundational concepts of complex systems and network science in a brief, straightforward manner. It argues that the interdisciplinary appeal and trendiness of complexity in the social sciences has resulted in a bit of a morass of ambiguous terminology, internal inconsistencies, and overloaded concepts open to multiple interpretations. This chapter provides explanatory examples of these concepts familiar to scholars and practitioners not already versed in the technical science of complexity. Most relevant to this present study, this chapter expounds the theory of networks and network analysis that form the foundation of the remaining chapters.

Chapter 4 unpacks the connections between the built form and measures of its complexity. It contributes a new typology of tools and metrics from different scientific disciplines to assess measures of complexity that apply to urban form and particularly to urban design’s scale of intervention. In particular, the measures of network structure characterize the complexity of the circulation network in terms of density, permeability, resilience, and connectedness. These attributes influence the way an urban system’s physical links can structure complex interactions, connections, and dynamics. The analytical framework developed here is generalizable to empirical research of multiple neighborhood types and design standards.

Chapter 5 offers the primary contribution of this dissertation – OSMnx, a new tool to download, construct, correct, analyze, and visualize urban street networks using OpenStreetMap data. Street network analysis in the urban planning literature suffers from challenges of usability, planarity, reproducibility, and sample sizes. To address these challenges, the primary thrust of this study was to create a tool, OSMnx, to make the collection of data and creation and analysis of street networks simple, consistent, and automatable. OSMnx contributes five significant new capabilities for researchers: first, the automatic downloading of administrative place boundaries and shapefiles; second, the tailored and automated downloading and constructing of street networks from OpenStreetMap; third, the automatic correction and simplification of network topology; fourth, the ability to save street networks to disk as shapefiles, GraphML, or SVG files; and fifth, the ability to analyze street networks, calculate routes, visualize the networks, and calculate network metrics and statistics. These metrics and statistics include both those common in urban design and transportation studies, and metrics that measure the complexity of the network.

OSMnx allows researchers and planners to download spatial geometries and construct, project, visualize, and analyze complex street networks. It makes the acquisition, construction, and analysis of urban street networks easy, consistent, and reproducible for powerful and consistent research, transportation engineering, and urban design. OSMnx is built on top of NetworkX, matplotlib, and geopandas for rich network analytic capabilities, beautiful and simple visualizations, and fast spatial queries with r-tree indexing. Finally, chapter 5 adapts measures from traditional network analysis to make them better-suited to accurately describing the physical form of street intersections and network connectivity – namely, by adapting abstract nodes and directed edges into faithful representations of intersections, dead-ends, and physical streets.

In turn, chapter 6 presents a small case study to simply but plainly demonstrate the usage of OSMnx for research. It collects three small half-kilometer sections of the street network in different neighborhoods in Portland, Oregon to perform a cross-sectional analysis. This scale of analysis and sample size are small, but they provide straightforward examples to tie-together the network concepts presented in chapter 3, the network measures presented in chapter 4, and the methodological tool presented in chapter 5. This chapter thus serves to knit these preceding threads together empirically.

This chapter also presents empirical findings of three street network sections in Portland, Oregon and uses the quantitative measures to compare and contrast these network sections. It introduces these neighborhoods first from a brief qualitative and historical perspective, then explores their comparative quantitative measures of network complexity and structure. The findings identifies significant chokepoints in the suburban network and demonstrates how there could be substantial gains in network robustness if one-way streets in the dense, orthogonal downtown were converted to two-way streets.

Following up from this small case study, chapter 7 presents a large multi-scale analysis of 27,000 street networks. The empirical literature on street network analysis is growing ever richer, but suffers from some limitations. First, sample sizes tend to be fairly small due to data availability, gathering, and processing constraints. Second, reproducibility is difficult when the dozens of decisions that go into analysis – such as spatial extents, topological simplification and correction, definitions of nodes and edges, etc. – are ad hoc or only partly reported. Third, and related to the first two, studies frequently oversimplify to planar or undirected primal graphs for tractability, or use dual graphs despite the loss of geographic and metric information. Fourth, the current landscape of tools and methods offers no ideal technique that balances usability, customizability, reproducibility, and scalability in acquiring, constructing, and analyzing network data.

This fourth limitation above was addressed by introducing OSMnx and demonstrating its use in a small case study of Portland, Oregon in chapters 5 and 6. Chapter 7 addressed the first three limitations by conducting an analysis of street networks at multiple scales, with large sample sizes, with clearly defined network definitions and extents for reproducibility, and using non-planar, directed graphs. In particular, it examines urban street networks – represented as primal, non-planar, weighted multidigraphs with possible self-loops – through the framework of complexity developed in this dissertation, focusing on structure, density, permeability, connectedness, centrality, and resilience.

Most studies in the street network literature that conduct topological and/or metric analysis tend to have sample sizes ranging around 5 to 50 networks. This chapter instead conducted a large analysis of over 27,000 urban street networks at multiple overlapping scales across the United States. Namely, it examined the street networks of every U.S. incorporated city and town, urbanized area, and Zillow-defined neighborhood. In total, the study presented in this chapter uses OSMnx to download, construct, and analyze 497 urbanized areas’ street networks, 19,655 cities’ and towns’ street networks, and 6,857 neighborhoods’ street networks. It uses these street networks to conduct four analyses: at the metropolitan scale, at the municipal scale, at the neighborhood scale, and a case study looking deeper at the neighborhood-scale street networks in the city of San Francisco.

An adaptation of chapter 2 is currently in press to be published as a journal article. Chapters 5 and 6 are currently in the final stages of consolidation and preparation to be submitted as a journal article. Chapters 4 and 7 are individually in the beginning stages of the article preparation process.

### Contribution to Planning Practice

This dissertation makes methodological and empirical contributions to urban planning practice.

## Future Research

Future research includes developing a decision-tree algorithm to impute street width from various attributes of the edge. Figure 5.9 used OpenStreetMap’s highway type attribute in conjunction with an ad hoc approximation of local street widths in each geography to map from highway type to an inferred width in meters to a width in pixels given the dimensions of the image raster. OpenStreetMap also has attributes for street width (though this is infrequently non-null) and number of lanes (though lane widths vary from place to place. A better future algorithm might first check street width; if it is not present it would estimate the width from the number of lanes; if this too is not present, it would fall back on some default value for the highway type.

Bibliography

Agryzkov, T., Oliver, J. L., Tortosa, L., & Vicent, J. F. (2012). An algorithm for ranking the nodes of an urban network based on the concept of PageRank vector. *Applied Mathematics and Computation*, *219*(4), 2186–2193. https://doi.org/10.1016/j.amc.2012.08.064

Agryzkov, T., Oliver, J. L., Tortosa, L., & Vicent, J. F. (2013). A Model to Visualize Information in a Complex Streets’ Network. In S. Omatu, J. Neves, J. M. C. Rodriguez, J. F. Paz Santana, & S. R. Gonzalez (Eds.), *Distributed Computing and Artificial Intelligence* (Vol. 217, pp. 129–136). Cham: Springer International Publishing. Retrieved from http://link.springer.com/10.1007/978-3-319-00551-5\_16

Albert, R., & Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Reviews of Modern Physics*, *74*(1), 47.

Alexander, C. (1964). *Notes on the Synthesis of Form*. Cambridge: Harvard University Press.

Alexander, C. (1965). A City Is Not a Tree. *Architectural Forum*, *122*, 58–62.

Aligica, P. D. (2014). *Institutional Diversity and Political Economy: The Ostroms and Beyond*. Oxford University Press.

Allen, P. (2012). Cities: The Visible Expression of Co-evolving Complexity. In J. Portugali, H. Meyer, E. Stolk, & E. Tan (Eds.), *Complexity Theories of Cities Have Come of Age* (pp. 67–89). Berlin, Heidelberg: Springer Berlin Heidelberg.

Alpigini, J. J. (2004). Dynamical System Visualization and Analysis via Performance Maps. *Information Visualization*, *3*(4), 271–287. https://doi.org/10.1057/palgrave.ivs.9500082

Argote-Cabanero, J., Daganzo, C. F., & Lynn, J. W. (2015). Dynamic control of complex transit systems. *Transportation Research Part B: Methodological*, *81*, 146–160. https://doi.org/10.1016/j.trb.2015.09.003

Aziz-Alaoui, M., & Bertelle, C. (2009). *From System Complexity to Emergent Properties*. Springer Science & Business Media.

Babbs, C. F. (2014). Initiation of Ventricular Fibrillation by a Single Ectopic Beat in Three Dimensional Numerical Models of Ischemic Heart Disease: Abrupt Transition to Chaos. *Journal of Clinical & Experimental Cardiology*, *5*(10), 2–11. https://doi.org/10.4172/2155-9880.1000346

badger-debunking-cul-de-sac.pdf. (n.d.).

Bak, P., Tang, C., & Wiesenfeld, K. (1988). Self-organized criticality. *Physical Review A*, *38*(1), 364.

Baran, P. K., Rodríguez, D. A., & Khattak, A. J. (2008). Space Syntax and Walking in a New Urbanist and Suburban Neighbourhoods. *Journal of Urban Design*, *13*(1), 5–28. https://doi.org/10.1080/13574800701803498

Barrat, A., Barthelemy, M., Pastor-Satorras, R., & Vespignani, A. (2004). The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(11), 3747–3752.

Barrington-Leigh, C., & Millard-Ball, A. (2015a). A century of sprawl in the United States. *Proceedings of the National Academy of Sciences*, *112*(27), 8244–8249.

Barrington-Leigh, C., & Millard-Ball, A. (2015b). A century of sprawl in the United States. *Proceedings of the National Academy of Sciences*, *112*(27), 8244–8249.

Barthelemy, M. (2004). Betweenness centrality in large complex networks. *The European Physical Journal B - Condensed Matter*, *38*(2), 163–168. https://doi.org/10.1140/epjb/e2004-00111-4

Barthelemy, M. (2011). Spatial networks. *Physics Reports*, *499*(1–3), 1–101. https://doi.org/10.1016/j.physrep.2010.11.002

Barthelemy, M., Bordin, P., Berestycki, H., & Gribaudi, M. (2013). Self-organization versus top-down planning in the evolution of a city. *Scientific Reports*, *3*. https://doi.org/10.1038/srep02153

Barthélemy, M., & Flammini, A. (2008). Modeling Urban Street Patterns. *Physical Review Letters*, *100*(13). https://doi.org/10.1103/PhysRevLett.100.138702

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

Batty, M. (2013). *The New Science of Cities*. Cambridge, MA: The MIT Press.

Batty, M., & Longley, P. (1994). *Fractal Cities: A Geometry of Form and Function*. London: Academic Press.

Batty, M., & Marshall, S. (2012). The Origins of Complexity Theory in Cities and Planning. In J. Portugali, H. Meyer, E. Stolk, & E. Tan (Eds.), *Complexity Theories of Cities Have Come of Age* (pp. 21–45). Berlin: Springer.

Batty, M., & Xie, Y. (1999). Self-Organized Criticality and Urban Development. *Discrete Dynamics in Nature and Society*, *3*(2–3), 109–124.

Beineke, L. W., Oellermann, O. R., & Pippert, R. E. (2002). The average connectivity of a graph. *Discrete Mathematics*, *252*(1), 31–45.

Benguigui, L., Czamanski, D., Marinov, M., & Portugali, Y. (2000). When and Where Is a City Fractal? *Environment and Planning B*, *27*(4), 507–519. https://doi.org/10.1068/b2617

Bertuglia, C. S., Bianchi, G., & Mela, A. (Eds.). (1998). *The city and its sciences*. Heidelberg ; New York: Physica-Verlag.

Beven, K., & Freer, J. (2001). Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *Journal of Hydrology*, *249*(1), 11–29.

Boarnet, M. G., & Crane, R. (2001). *Travel by design: the influence of urban form on travel*. Oxford; New York: Oxford University Press. Retrieved from http://site.ebrary.com/id/10086949

Bonchev, D., & Buck, G. A. (2005). Quantitative Measures of Network Complexity. In D. Bonchev & D. H. Rouvray (Eds.), *Complexity in Chemistry, Biology, and Ecology* (pp. 191–235). Boston, MA: Springer US. Retrieved from http://link.springer.com/10.1007/0-387-25871-X\_5

Bordoloi, R., Mote, A., Sarkar, P. P., & Mallikarjuna, C. (2013). Quantification of Land Use Diversity in The Context of Mixed Land Use. *Procedia - Social and Behavioral Sciences*, *104*, 563–572. https://doi.org/10.1016/j.sbspro.2013.11.150

Bourdic, L., Salat, S., & Nowacki, C. (2012). Assessing cities: a new system of cross-scale spatial indicators. *Building Research & Information*, *40*(5), 592–605. https://doi.org/10.1080/09613218.2012.703488

Bradley, E. (2003). Time Series Analysis. In D. Hand & M. Berthold (Eds.), *Intelligent Data Analysis: An Introduction* (2nd ed.). Berlin: Springer-Verlag.

Bradley, E., & Kantz, H. (2015). Nonlinear Time-Series Analysis Revisited. *Chaos*, *25*(9), 97610. https://doi.org/10.1063/1.4917289

Brandes, U., & Erlebach, T. (Eds.). (2005). *Network analysis: methodological foundations*. Berlin ; New York: Springer.

Brown, T. A. (1996). Measuring Chaos Using the Lyapunov Exponent. In L. D. Kiel & E. Elliott (Eds.), *Chaos Theory in the Social Sciences* (pp. 53–66). Ann Arbor: University of Michigan Press.

Byrne, D. (2001). What is complexity science? Thinking as a realist about measurement and cities and arguing for natural history. *Emergence, A Journal of Complexity Issues in Organizations and Management*, *3*(1), 61–76.

Cao, J., & Menendez, M. (2015). System dynamics of urban traffic based on its parking-related-states. *Transportation Research Part B: Methodological*. https://doi.org/10.1016/j.trb.2015.07.018

Carlson, C., Aytur, S., Gardner, K., & Rogers, S. (2012). Complexity in Built Environment, Health, and Destination Walking: A Neighborhood-Scale Analysis. *Journal of Urban Health*, *89*(2), 270–284. https://doi.org/10.1007/s11524-011-9652-8

Cartwright, T. J. (1991). Planning and Chaos Theory. *Journal of the American Planning Association*, *57*(1), 44–56. https://doi.org/10.1080/01944369108975471

Cavalcante, A., Mansouri, A., Kacha, L., Barros, A. K., Takeuchi, Y., Matsumoto, N., & Ohnishi, N. (2014). Measuring Streetscape Complexity Based on the Statistics of Local Contrast and Spatial Frequency. *PLoS ONE*, *9*(2), e87097. https://doi.org/10.1371/journal.pone.0087097

Cervero, R., & Kockelman, K. (1997). Travel demand and the 3Ds: density, diversity, and design. *Transportation Research Part D: Transport and Environment*, *2*(3), 199–219.

Chan, K.-S., & Tong, H. (2013). *Chaos: A Statistical Perspective*. New York: Springer Science & Business Media.

Chen, C. (2006). *Information Visualization: Beyond the Horizon* (2nd ed.). London: Springer-Verlag.

Chen, W.-C. (2008). Nonlinear Dynamics and Chaos in a Fractional-Order Financial System. *Chaos, Solitons & Fractals*, *36*(5), 1305–1314. https://doi.org/10.1016/j.chaos.2006.07.051

Chen, Y., & Zhou, Y. (2008). Scaling Laws and Indications of Self-Organized Criticality in Urban Systems. *Chaos, Solitons & Fractals*, *35*(1), 85–98. https://doi.org/10.1016/j.chaos.2006.05.018

Chettiparamb, A. (2006). Metaphors in Complexity Theory and Planning. *Planning Theory*, *5*(1), 71–91. https://doi.org/10.1177/1473095206061022

Chin, W.-C.-B., & Wen, T.-H. (2015). Geographically Modified PageRank Algorithms: Identifying the Spatial Concentration of Human Movement in a Geospatial Network. *PLOS ONE*, *10*(10), e0139509. https://doi.org/10.1371/journal.pone.0139509

Clarke, K. C. (1986). Computation of the Fractal Dimension of Topographic Surfaces Using the Triangular Prism Surface Area Method. *Computers & Geosciences*, *12*(5), 713–722.

Clifton, K., Ewing, R., Knaap, G., & Song, Y. (2008). Quantitative analysis of urban form: a multidisciplinary review. *Journal of Urbanism: International Research on Placemaking and Urban Sustainability*, *1*(1), 17–45. https://doi.org/10.1080/17549170801903496

Corcoran, P., Mooney, P., & Bertolotto, M. (2013). Analysing the growth of OpenStreetMap networks. *Spatial Statistics*, *3*, 21–32. https://doi.org/10.1016/j.spasta.2013.01.002

Costa, L. da F., Rodrigues, F. A., Travieso, G., & Villas Boas, P. R. (2007). Characterization of complex networks: A survey of measurements. *Advances in Physics*, *56*(1), 167–242. https://doi.org/10.1080/00018730601170527

Cranmer, S. J., Leifeld, P., McClurg, S. D., & Rolfe, M. (2016). Navigating the range of statistical tools for inferential network analysis. *American Journal of Political Science*. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/ajps.12263/full

Crucitti, P., Latora, V., & Porta, S. (2006a). Centrality in networks of urban streets. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *16*(1), 15113. https://doi.org/10.1063/1.2150162

Crucitti, P., Latora, V., & Porta, S. (2006b). Centrality Measures in Spatial Networks of Urban Streets. *Physical Review E*, *73*.

Crucitti, P., Latora, V., & Porta, S. (2006c). Centrality measures in spatial networks of urban streets. *Physical Review E*, *73*(3). https://doi.org/10.1103/PhysRevE.73.036125

Danforth, C. M. (2013, March 17). Chaos in an Atmosphere Hanging on a Wall. Retrieved from http://mpe2013.org/2013/03/17/chaos-in-an-atmosphere-hanging-on-a-wall/

Deppisch, S., & Schaerffer, M. (2011). Given the Complexity of Large Cities, Can Urban Resilience be Attained at All? In B. Müller (Ed.), *German Annual of Spatial Research and Policy 2010* (pp. 25–33). Berlin, Heidelberg: Springer Berlin Heidelberg. Retrieved from http://link.springer.com/10.1007/978-3-642-12785-4\_3

Dill, J. (2004). Measuring network connectivity for bicycling and walking. In *83rd Annual Meeting of the Transportation Research Board, Washington, DC* (pp. 11–15). Retrieved from http://reconnectingamerica.org/assets/Uploads/TRB2004-001550.pdf

Dingwell, J. B. (2006). Lyapunov Exponents. In M. Akay (Ed.), *Wiley Encyclopedia of Biomedical Engineering*. Hoboken: John Wiley & Sons.

Dorogovtsev-evolution-networks.pdf. (n.d.).

Drinan, J. M. (2015). Response to Talen et al.: Igniting the Dialogue: What Makes a Neighborhood Great? *Journal of the American Planning Association*, *81*(2), 141–142. https://doi.org/10.1080/01944363.2015.1068652

Dunkel, A. (2015). Visualizing the perceived environment using crowdsourced photo geodata. *Landscape and Urban Planning*, *142*, 173–186. https://doi.org/10.1016/j.landurbplan.2015.02.022

Eikelboom, T., Janssen, R., & Stewart, T. J. (2015). A spatial optimization algorithm for geodesign. *Landscape and Urban Planning*, *144*, 10–21.

Elsheshtawy, Y. (1997). Urban Complexity: Toward the Measurement of the Physical Complexity of Street-Scapes. *Journal of Architectural and Planning Research*, *14*(4), 301–316.

Ermagun, A. (2016). An Introduction to the Network Weight Matrix. Retrieved from http://conservancy.umn.edu/handle/11299/181543

Evans, P. (2002). *Livable Cities: Urban Struggles for Livelihood and Sustainability*. Berkeley: University of California Press.

Ewing, R., & Cervero, R. (2010). Travel and the Built Environment: A Meta-Analysis. *Journal of the American Planning Association*, *76*(3), 265–294. https://doi.org/10.1080/01944361003766766

Ewing, R. H., & Clemente, O. (2013). *Measuring urban design: metrics for livable places*. Island Press.

Farmer, J. D., Ott, E., & Yorke, J. A. (1983). The Dimension of Chaotic Attractors. *Physica D: Nonlinear Phenomena*, 153–180.

fecht-sciam-grid-unlocked.pdf. (n.d.).

Feigenbaum, M. J. (1978). Quantitative Universality for a Class of Nonlinear Transformations. *Journal of Statistical Physics*, *19*(1), 25–52.

Feigenbaum, M. J. (1983). Universal Behavior in Nonlinear Systems. *Physica D: Nonlinear Phenomena*, *7*(1), 16–39.

Fischer-GIS-network-analysis.pdf. (n.d.).

Foti, F. (2014). *Behavioral Framework for Measuring Walkability and its Impact on Home Values and Residential Location Choices*. University of California, Berkeley, Berkeley, California.

Frame, M., Mandelbrot, B., & Neger, N. (2015). *Fractal Geometry*. Yale University. Retrieved from https://classes.yale.edu/fractals/

Frizzelle-tigerline-road-accuracy.pdf. (n.d.).

Garrick, N., & Marshall, W. (2009). The Effect of Street Network Design on Walking and Biking.

Gershenson, C. (2004, October 7). *Introduction to Chaos in Deterministic Systems*. University of Sussex. Retrieved from http://arxiv.org/abs/nlin/0308023

Gershenson, C., & Fernández, N. (2012). Complexity and information: Measuring emergence, self-organization, and homeostasis at multiple scales. *Complexity*, *18*(2), 29–44. https://doi.org/10.1002/cplx.21424

Giacomin, D. J., & Levinson, D. M. (2015). Road network circuity in metropolitan areas. *Environment and Planning B: Planning and Design*, *0*(0), 0–0. https://doi.org/10.1068/b130131p

Glass, L. (2009). Introduction to Controversial Topics in Nonlinear Science: Is the Normal Heart Rate Chaotic? *Chaos*, *19*(2), 28501. https://doi.org/10.1063/1.3156832

Gleich, D. F. (2015). PageRank Beyond the Web. *SIAM Review*, *57*(3), 321–363. https://doi.org/10.1137/140976649

Gleick, J. (1991). *Chaos: Making a New Science*. Indianapolis: Cardinal.

Goodchild, M. F. (2007). Citizens as sensors: the world of volunteered geography. *GeoJournal*, *69*(4), 211–221. https://doi.org/10.1007/s10708-007-9111-y

Grassberger, P., & Procaccia, I. (1983). Characterization of Strange Attractors. *Physical Review Letters*, *50*(5), 346–349.

Grebogi, C., Ott, E., & Yorke, J. A. (1987). Chaos, Strange Attractors, and Fractal Basin Boundaries in Nonlinear Dynamics. *Science*, *238*(4827), 632–638.

Guastello, S. J. (2013). *Chaos, Catastrophe, and Human Affairs: Applications of Nonlinear Dynamics to Work, Organizations, and Social Evolution*. New York: Psychology Press.

Gudmundsson, A., & Mohajeri, N. (2013). Entropy and order in urban street networks. *Scientific Reports*, *3*. https://doi.org/10.1038/srep03324

Guégan, D. (2009). Chaos in Economics and Finance. *Annual Reviews in Control*, *33*(1), 89–93. https://doi.org/10.1016/j.arcontrol.2009.01.002

Hagberg, A., & Conway, D. (2010). Hacking social networks using the Python programming language. Presented at the Sunbelt 2010: International Network for Social Network Analysis.

Hajrasouliha, A., & Yin, L. (2015a). The impact of street network connectivity on pedestrian volume. *Urban Studies*, *52*(13), 2483–2497. https://doi.org/10.1177/0042098014544763

Hajrasouliha, A., & Yin, L. (2015b). The impact of street network connectivity on pedestrian volume. *Urban Studies*, *52*(13), 2483–2497. https://doi.org/10.1177/0042098014544763

Haken, H. (2012). Complexity and Complexity Theories: Do These Concepts Make Sense? In J. Portugali, H. Meyer, E. Stolk, & E. Tan (Eds.), *Complexity Theories of Cities Have Come of Age* (pp. 7–20). Berlin, Heidelberg: Springer Berlin Heidelberg.

Hall, P. (1996). *Cities of Tomorrow: An Intellectual History of Urban Planning and Design in the Twentieth Century* (2nd ed.). Malden, MA: Blackwell Publishers.

hamblin-fat-in-suburbs.pdf. (n.d.).

Hamouche, M. B. (2009). Can Chaos Theory Explain Complexity In Urban Fabric? Applications in Traditional Muslim Settlements. *Nexus Network Journal: Architecture and Mathematics*, *11*(2), 217–242. https://doi.org/10.1007/s00004-008-0088-8

handy-community-design-active-travel.pdf. (n.d.).

Harvey, D. (2010). *Social Justice and the City*. Athens: University of Georgia Press.

Hastings, A., Hom, C. L., Ellner, S., Turchin, P., & Godfray, H. C. J. (1993). Chaos in Ecology: Is Mother Nature a Strange Attractor? *Annual Review of Ecology and Systematics*, *24*, 1–33.

Hénon, M. (1976). A Two-Dimensional Mapping with a Strange Attractor. *Communications in Mathematical Physics*, *50*(1), 69–77.

Hillier, B., Leaman, A., Stansall, P., & Bedford, M. (1976). Space Syntax. *Environment and Planning B*, *3*, 147–185.

Hoch, C., Zellner, M., Milz, D., Radinsky, J., & Lyons, L. (2015). Seeing is not believing: cognitive bias and modelling in collaborative planning. *Planning Theory & Practice*, *16*(3), 319–335. https://doi.org/10.1080/14649357.2015.1045015

Hong, Z., & Dong, J. (2010). Chaos Theory and Its Application in Modern Cryptography (Vol. 7, pp. 332–334). Presented at the International Conference on Computer Application and System Modeling (ICCASM 2010), Taiyuan, China.

Hoshi, R. A., Pastre, C. M., Vanderlei, L. C. M., & Godoy, M. F. (2013). Poincaré Plot Indexes of Heart Rate Variability: Relationships with Other Nonlinear Variables. *Autonomic Neuroscience*, *177*(2), 271–274. https://doi.org/10.1016/j.autneu.2013.05.004

Hu, Y., Wu, Q., & Zhu, D. (2008). Topological patterns of spatial urban street networks. In *2008 4th International Conference on Wireless Communications, Networking and Mobile Computing* (pp. 1–4). IEEE. Retrieved from http://ieeexplore.ieee.org/xpls/abs\_all.jsp?arnumber=4681175

Huang, X., Zhao, Y., Ma, C., Yang, J., Ye, X., & Zhang, C. (2016). TrajGraph: A Graph-Based Visual Analytics Approach to Studying Urban Network Centralities Using Taxi Trajectory Data. *IEEE Transactions on Visualization and Computer Graphics*, *22*(1), 160–169. https://doi.org/10.1109/TVCG.2015.2467771

Huikuri, H. V., Mäkikallio, T. H., Peng, C.-K., Goldberger, A. L., Hintze, U., Møller, M., & others. (2000). Fractal Correlation Properties of RR Interval Dynamics and Mortality in Patients with Depressed Left Ventricular Function after an Acute Myocardial Infarction. *Circulation*, *101*(1), 47–53.

Hunt, B. R., & Ott, E. (2015). Defining Chaos. *Chaos*, *25*(9), 97618. https://doi.org/10.1063/1.4922973

Innes, J. E., & Booher, D. E. (2010). *Planning with Complexity*. London: Routledge.

Jabareen, Y. (2013). Planning the resilient city: Concepts and strategies for coping with climate change and environmental risk. *Cities*, *31*, 220–229. https://doi.org/10.1016/j.cities.2012.05.004

Jackson, K. T. (1985). *Crabgrass Frontier: The Suburbanization of the United States*. New York: Oxford University Press.

Jacobs, A. B. (1995). *Great Streets*. Cambridge, MA: The MIT Press.

Jacobs, J. (1961). *The Death and Life of Great American Cities* (1992 Edition). New York: Vintage Books.

Jguirim, I., Brosset, D., & Claramunt, C. (2014). Functional and Structural Analysis of an Urban Space Extended from Space Syntax. GIScience. Retrieved from https://www.researchgate.net/profile/David\_Brosset/publication/278049002\_Functional\_and\_Structural\_Analysis\_of\_an\_Urban\_Space\_Extended\_from\_Space\_Syntax/links/557bd90108aeb61eae21d00b.pdf

Jiang, B. (2009). Ranking spaces for predicting human movement in an urban environment. *International Journal of Geographical Information Science*, *23*(7), 823–837.

Jiang, B., & Claramunt, C. (2002). Integration of space syntax into GIS: new perspectives for urban morphology. *Transactions in GIS*, *6*(3), 295–309.

Jiang, B., & Claramunt, C. (2004a). Topological analysis of urban street networks. *Environment and Planning B: Planning and Design*, *31*(1), 151–162. https://doi.org/10.1068/b306

Jiang, B., & Claramunt, C. (2004b). Topological Analysis of Urban Street Networks. *Environment and Planning B: Planning and Design*, *31*(1), 151–162. https://doi.org/10.1068/b306

Jiang, B., Duan, Y., Lu, F., Yang, T., & Zhao, J. (2014). Topological structure of urban street networks from the perspective of degree correlations. *Environment and Planning B: Planning and Design*, *41*(5), 813–828.

Jiang, B., Yin, J., & Zhao, S. (2009). Characterizing the human mobility pattern in a large street network. *Physical Review E*, *80*(2), 21136.

Jokar Arsanjani, J., Zipf, A., Mooney, P., & Helbich, M. (Eds.). (2015a). *OpenStreetMap in GIScience*. Cham: Springer International Publishing. Retrieved from http://link.springer.com/10.1007/978-3-319-14280-7

Jokar Arsanjani, J., Zipf, A., Mooney, P., & Helbich, M. (Eds.). (2015b). *OpenStreetMap in GIScience*. Cham: Springer International Publishing. Retrieved from http://link.springer.com/10.1007/978-3-319-14280-7

Kantz, H., Radons, G., & Yang, H. (2013). The Problem of Spurious Lyapunov Exponents in Time Series Analysis and Its Solution by Covariant Lyapunov Vectors. *Journal of Physics A: Mathematical and Theoretical*, *46*(25), 254009. https://doi.org/10.1088/1751-8113/46/25/254009

Karduni, A., Kermanshah, A., & Derrible, S. (2016). A protocol to convert spatial polyline data to network formats and applications to world urban road networks. *Scientific Data*, *3*, 160046. https://doi.org/10.1038/sdata.2016.46

Karimi, K. (2012). A configurational approach to analytical urban design:“Space syntax”methodology. *Urban Design International*, *17*(4), 297–318.

Kekre, H. B., Sarode, T., & Halarnkar, P. N. (2014). A Study of Period Doubling in Logistic Map for Shift Parameter. *International Journal of Engineering Trends and Technology*, *13*(6), 281–286.

Knight, P. L., & Marshall, W. E. (2015). The metrics of street network connectivity: their inconsistencies. *Journal of Urbanism: International Research on Placemaking and Urban Sustainability*, *8*(3), 241–259. https://doi.org/10.1080/17549175.2014.909515

Krasny, M. E., Russ, A., Tidball, K. G., & Elmqvist, T. (2014). Civic ecology practices: Participatory approaches to generating and measuring ecosystem services in cities. *Ecosystem Services*, *7*, 177–186. https://doi.org/10.1016/j.ecoser.2013.11.002

Krizek, K., & Waddell, P. (2002). Analysis of lifestyle choices: Neighborhood type, travel patterns, and activity participation. *Transportation Research Record: Journal of the Transportation Research Board*, (1807), 119–128.

Larice, M., & Macdonald, E. (2007). *The Urban Design Reader* (1st ed.). New York: Routledge.

Layek, G. C. (2015). *An Introduction to Dynamical Systems and Chaos*. New Delhi: Springer India.

Levinson, D. (2012). Network Structure and City Size. *PLoS ONE*, *7*(1), e29721. https://doi.org/10.1371/journal.pone.0029721

Levinson, D., & El-Geneidy, A. (2009). The minimum circuity frontier and the journey to work. *Regional Science and Urban Economics*, *39*(6), 732–738. https://doi.org/10.1016/j.regsciurbeco.2009.07.003

Levinson, D., & Huang, A. (2012). A Positive Theory of Network Connectivity. *Environment and Planning B: Planning and Design*, *39*(2), 308–325. https://doi.org/10.1068/b37094

Li, T.-Y., & Yorke, J. A. (1975). Period Three Implies Chaos. *The American Mathematical Monthly*, *82*(10), 985–992. https://doi.org/10.2307/2318254

Li, W., Wang, K., & Su, H. (2011). Optimal Harvesting Policy for Stochastic Logistic Population Model. *Applied Mathematics and Computation*, *218*(1), 157–162. https://doi.org/10.1016/j.amc.2011.05.079

Liu, Y.-Y., Slotine, J.-J., & Barabási, A.-L. (2011). Controllability of complex networks. *Nature*, *473*(7346), 167–173. https://doi.org/10.1038/nature10011

Lloyd, S. (2001). Measures of complexity: a nonexhaustive list. *IEEE Control Systems Magazine*, *21*(4), 7–8.

Lorenz, E. N. (1963). Deterministic Nonperiodic Flow. *Journal of the Atmospheric Sciences*, *20*, 130–141.

Makris, G., & Antoniou, I. (2012). Cryptography with Chaos. In *Proceedings of the 5th Chaotic Modeling and Simulation International Conference* (pp. 309–318). Athens, Greece.

Malamud, B., Morein, G., & Turcotte, D. (1998). Forest Fires: An Example of Self-Organized Critical Behavior. *Science*, *281*(5384), 1840–1842.

Mandelbrot, B. B. (1967). How Long Is the Coast of Britain? *Science*, *156*(3775), 636–638.

Mandelbrot, B. B. (1983). *The Fractal Geometry of Nature*. New York: Macmillan.

Mandelbrot, B. B. (1999). *Multifractals and 1/f Noise*. New York: Springer.

Manson, S. M. (2001). Simplifying complexity: a review of complexity theory. *Geoforum*, *32*, 405–414.

Manson, S., & O’Sullivan, D. (2006). Complexity theory in the study of space and place. *Environment and Planning A*, *38*(4), 677–692. https://doi.org/10.1068/a37100

Marcus, L., & Legeby, A. (2012). The need for co-presence in urban complexity: Measuring social capital using space syntax. In *Eigth International Space Syntax Symposium*. Retrieved from http://www.diva-portal.org/smash/record.jsf?pid=diva2:470211

Marshall, S. (2012). Planning, Design and the Complexity of Cities. In J. Portugali, H. Meyer, E. Stolk, & E. Tan (Eds.), *Complexity Theories of Cities Have Come of Age* (pp. 191–205). Berlin, Heidelberg: Springer Berlin Heidelberg.

Marshall, W. E., & Garrick, N. W. (2010). Street network types and road safety: A study of 24 California cities. *Urban Design International*, *15*(3), 133–147.

Marshall, W. E., Piatkowski, D. P., & Garrick, N. W. (2014). Community design, street networks, and public health. *Journal of Transport & Health*, *1*(4), 326–340. https://doi.org/10.1016/j.jth.2014.06.002

marshall-street-networks-chapter.pdf. (n.d.).

Mattsson, L.-G., & Jenelius, E. (2015). Vulnerability and resilience of transport systems – A discussion of recent research. *Transportation Research Part A: Policy and Practice*, *81*, 16–34. https://doi.org/10.1016/j.tra.2015.06.002

May, R. M. (1974). Biological Populations with Nonoverlapping Generations: Stable Points, Stable Cycles, and Chaos. *Science*, *186*(4164), 645–647.

May, R. M. (1976). Simple Mathematical Models with Very Complicated Dynamics. *Nature*, *261*, 459–467.

Mitchell, M. (2009). *Complexity: A Guided Tour* (Later prt.). Oxford University Press, USA.

Mitzenmacher, M. (2004). A Brief History of Generative Models for Power Law and Lognormal Distributions. *Internet Mathematics*, *1*(2).

Moroni, S. (2010). Rethinking the theory and practice of land-use regulation: Towards nomocracy. *Planning Theory*, *9*(2), 137–155.

Moroni, S. (2015). Complexity and the inherent limits of explanation and prediction: Urban codes for self-organising cities. *Planning Theory*, *14*(3), 248–267.

Mumford, L. (1961). *The City in History: Its Origins, Its Transformations, and Its Prospects*. San Diego: Harcourt Brace Jovanovich.

Murcio, R., Morphet, R., Gershenson, C., & Batty, M. (2015). Urban transfer entropy across scales. *arXiv Preprint arXiv:1505.02761*. Retrieved from http://arxiv.org/abs/1505.02761

Newman, M. (2010). *Networks: An Introduction* (1 edition). Oxford ; New York: Oxford University Press.

Newman, M. E. (2003). The structure and function of complex networks. *SIAM Review*, *45*(2), 167–256.

Nicolis, G., & Prigogine, I. (1977). *Self-Organization in Nonequilibrium Systems: From Dissipative Structures to Order through Fluctuations* (1 edition). New York: Wiley.

Opsahl, T., & Panzarasa, P. (2009). Clustering in weighted networks. *Social Networks*, *31*(2), 155–163. https://doi.org/10.1016/j.socnet.2009.02.002

Orishimo, I. (1987). An approach to urban dynamics. *Geographical Analysis*, *19*(3), 200–210.

Ostwald, M. J. (2013). The Fractal Analysis of Architecture: Calibrating the Box-Counting Method Using Scaling Coefficient and Grid Disposition Variables. *Environment and Planning B*, *40*(4), 644–663. https://doi.org/10.1068/b38124

O’Sullivan, A. (2008). *Urban Economics* (7 edition). Boston: McGraw-Hill/Irwin.

O’Sullivan, D. (2014). Spatial Network Analysis. In M. M. Fischer & P. Nijkamp (Eds.), *Handbook of Regional Science* (pp. 1253–1273). Berlin, Heidelberg: Springer Berlin Heidelberg. Retrieved from http://link.springer.com/10.1007/978-3-642-23430-9\_67

O’Sullivan, D., & Manson, S. M. (2015). Do Physicists Have Geography Envy? And What Can Geographers Learn from It? *Annals of the Association of American Geographers*, *105*(4), 704–722. https://doi.org/10.1080/00045608.2015.1039105

O’Sullivan, D., & Perry, G. L. W. (2013). *Spatial simulation: exploring pattern and process*. Chichester, West Sussex, U.K.: John Wiley & Sons Inc. Retrieved from http://site.ebrary.com/id/10748659

OTREC-RR-11-04\_Final.pdf. (n.d.).

Over, M., Schilling, A., Neubauer, S., & Zipf, A. (2010). Generating web-based 3D City Models from OpenStreetMap: The current situation in Germany. *Computers, Environment and Urban Systems*, *34*(6), 496–507. https://doi.org/10.1016/j.compenvurbsys.2010.05.001

Oxley, L., & George, D. A. R. (2007). Economics on the Edge of Chaos: Some Pitfalls of Linearizing Complex Systems. *Environmental Modelling & Software*, *22*(5), 580–589. https://doi.org/10.1016/j.envsoft.2005.12.018

pacific-std-culdesac.pdf. (n.d.).

Packard, N. H., Crutchfield, J. P., Farmer, J. D., & Shaw, R. S. (1980). Geometry from a Time Series. *Physical Review Letters*, *45*(9), 712–716.

Parrott, L. (2010). Measuring ecological complexity. *Ecological Indicators*, *10*(6), 1069–1076. https://doi.org/10.1016/j.ecolind.2010.03.014

Parthasarathi, P., Hochmair, H., & Levinson, D. (2012). Network Structure and Spatial Separation. *Environment and Planning B: Planning and Design*, *39*(1), 137–154. https://doi.org/10.1068/b36139

Parthasarathi, P., Hochmair, H., & Levinson, D. (2015). Street network structure and household activity spaces. *Urban Studies*, *52*(6), 1090–1112. https://doi.org/10.1177/0042098014537956

Parthasarathi, P. K. (2011). *Network structure and travel*. University of Minnesota. Retrieved from http://conservancy.umn.edu/handle/11299/115927

Parthasarathi, P., Levinson, D., & Hochmair, H. (2013). Network Structure and Travel Time Perception. *PLoS ONE*, *8*(10), e77718. https://doi.org/10.1371/journal.pone.0077718

Pastijn, H. (2006). Chaotic Growth with the Logistic Model of P.F. Verhulst. In M. Ausloos & M. Dirickx (Eds.), *The Logistic Map and the Route to Chaos* (pp. 3–11). Berlin: Springer-Verlag.

Perry, C. (2007). The Neighborhood Unit. In M. Larice & E. Macdonald (Eds.), *The Urban Design Reader* (1st ed., pp. 54–65). New York: Routledge.

Porta, S., Crucitti, P., & Latora, V. (2006a). The network analysis of urban streets: A dual approach. *Physica A: Statistical Mechanics and Its Applications*, *369*(2), 853–866. https://doi.org/10.1016/j.physa.2005.12.063

Porta, S., Crucitti, P., & Latora, V. (2006b). The network analysis of urban streets: a primal approach. *Environment and Planning B: Planning and Design*, *33*(5), 705–725. https://doi.org/10.1068/b32045

Porta, S., Romice, O., Maxwell, J. A., Russell, P., & Baird, D. (2014). Alterations in scale: Patterns of change in main street networks across time and space. *Urban Studies*, *51*(16), 3383–3400. https://doi.org/10.1177/0042098013519833

Portugali, J. (2006). Complexity theory as a link between space and place. *Environment and Planning A*, *38*(4), 647–664. https://doi.org/10.1068/a37260

Portugali, J. (2012). Complexity theories of cities: Achievements, criticism and potentials. In *Complexity Theories of Cities Have Come of Age* (pp. 47–62). Springer.

Prigogine, I. (1997). *The End of Certainty* (1 edition). New York: Free Press.

Puu, T. (2013). *Attractors, Bifurcations, & Chaos: Nonlinear Phenomena in Economics* (2nd ed.). New York: Springer Science & Business Media.

RANKING SPACES FOR PREDICTING HUMAN MOVEMENT IN AN URBAN ENV.doc. (n.d.).

Ratti, C. (2004a). Space syntax: some inconsistencies. *Environment and Planning B: Planning and Design*, *31*(4), 487–499. https://doi.org/10.1068/b3019

Ratti, C. (2004b). Space syntax: some inconsistencies. *Environment and Planning B: Planning and Design*, *31*(4), 487–499. https://doi.org/10.1068/b3019

Ravulaparthy, S., & Goulias, K. (2014). Characterizing the Composition of Economic Activities in Central Locations: Graph-Theoretic Approach to Urban Network Analysis. *Transportation Research Record: Journal of the Transportation Research Board*, *2430*, 95–104. https://doi.org/10.3141/2430-10

reddit - osm project.pdf. (n.d.).

Reitsma, F. (2003). A response to simplifying complexity. *Geoforum*, *34*, 13–16.

Richards, D. (1996). From Individuals to Groups: The Aggregation of Votes and Chaotic Dynamics. In L. D. Kiel & E. Elliott (Eds.), *Chaos Theory in the Social Sciences* (pp. 89–116). Ann Arbor: University of Michigan Press.

Rickles, D., Hawe, P., & Shiell, A. (2007). A Simple Guide to Chaos and Complexity. *Journal of Epidemiology & Community Health*, *61*(11), 933–937. https://doi.org/10.1136/jech.2006.054254

Rodríguez, D. A., Khattak, A. J., & Evenson, K. R. (2006). Can New Urbanism Encourage Physical Activity?: Comparing a New Urbanist Neighborhood with Conventional Suburbs. *Journal of the American Planning Association*, *72*(1), 43–54. https://doi.org/10.1080/01944360608976723

Rosser, Jr., J. B. (1996). Chaos Theory and Rationality in Economics. In L. D. Kiel & E. Elliott (Eds.), *Chaos Theory in the Social Sciences* (pp. 199–213). Ann Arbor: University of Michigan Press.

Ruelle, D., & Takens, F. (1971). On the Nature of Turbulence. *Communications in Mathematical Physics*, *20*(3), 167–192.

Salat, S., Bourdic, L., & Nowacki, C. (2010). Assessing urban complexity. *International Journal of Sustainable Building Technology and Urban Development*, *1*(2), 160–167.

Salingaros, N. A. (2000). Complexity and Urban Coherence. *Journal of Urban Design*, *5*(3), 291–316. https://doi.org/10.1080/713683969

Salingaros, N. A. (2001). Fractals in the New Architecture. *Archimagazine*. Retrieved from http://zeta.math.utsa.edu/~yxk833/fractals.html

Sander, E., & Yorke, J. A. (2015). The Many Facets of Chaos. *International Journal of Bifurcation and Chaos*, *25*(4), 1530011. https://doi.org/10.1142/S0218127415300116

Schuster, P. (2015). Models: From exploration to prediction: Bad reputation of modeling in some disciplines results from nebulous goals. *Complexity*, 1–4. https://doi.org/10.1002/cplx.21729

Scott, J. C. (1999). *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed* (New edition). Yale University Press.

Sevtsuk, A. (2014). Location and Agglomeration The Distribution of Retail and Food Businesses in Dense Urban Environments. *Journal of Planning Education and Research*, *34*(4), 374–393.

Sevtsuk, A., & Mekonnen, M. (2012a). Urban network analysis. A new toolbox for ArcGIS. *Revue Internationale de Géomatique*, *22*(2), 287–305. https://doi.org/10.3166/rig.22.287-305

Sevtsuk, A., & Mekonnen, M. (2012b). Urban network analysis. A new toolbox for ArcGIS. *Revue Internationale de Géomatique*, *22*(2), 287–305. https://doi.org/10.3166/rig.22.287-305

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, *27*, 379–423, 623–656.

Shen, G. (2002a). Fractal Dimension and Fractal Growth of Urbanized Areas. *International Journal of Geographical Information Science*, *16*(5), 419–437. https://doi.org/10.1080/13658810210137013

Shen, G. (2002b). Fractal dimension and fractal growth of urbanized areas. *International Journal of Geographical Information Science*, *16*(5), 419–437. https://doi.org/10.1080/13658810210137013

Shilnikov, L. (2002). Bifurcations and Strange Attractors. *Proceedings of the International Congress of Mathematicians*, *III*, 349–372.

Shiner, J. S., Davison, M., & Landsberg, P. T. (1999). Simple measure for complexity. *Physical Review E*, *59*(2), 1459.

Singh, S. L., Mishra, S. N., & Sinkala, W. (2012). A New Iterative Approach to Fractal Models. *Communications in Nonlinear Science and Numerical Simulation*, *17*(2), 521–529. https://doi.org/10.1016/j.cnsns.2011.06.014

Siodla, J. (2015). Razing San Francisco: The 1906 disaster as a natural experiment in urban redevelopment. *Journal of Urban Economics*, *89*, 48–61. https://doi.org/10.1016/j.jue.2015.07.001

Smith, M. E. (2010). The archaeological study of neighborhoods and districts in ancient cities. *Journal of Anthropological Archaeology*, *29*(2), 137–154. https://doi.org/10.1016/j.jaa.2010.01.001

Song, Y., Merlin, L., & Rodriguez, D. (2013). Comparing measures of urban land use mix. *Computers, Environment and Urban Systems*, *42*, 1–13. https://doi.org/10.1016/j.compenvurbsys.2013.08.001

Song, Y., Popkin, B., & Gordon-Larsen, P. (2013). A national-level analysis of neighborhood form metrics. *Landscape and Urban Planning*, *116*, 73–85. https://doi.org/10.1016/j.landurbplan.2013.04.002

Sprott, J. C., & Xiong, A. (2015). Classifying and Quantifying Basins of Attraction. *Chaos*, *25*(8), 83101.

Stead, D., & Marshall, S. (2001). The relationships between urban form and travel patterns. An international review and evaluation. *European Journal of Transport and Infrastructure Research*, *1*(2), 113–141.

Stewart, I. (2000). The Lorenz Attractor Exists. *Nature*, *406*(6799), 948–949.

Stott, I., Soga, M., Inger, R., & Gaston, K. J. (2015). Land sparing is crucial for urban ecosystem services. *Frontiers in Ecology and the Environment*, *13*(7), 387–393. https://doi.org/10.1890/140286

Strano, E., Nicosia, V., Latora, V., Porta, S., & Barthélemy, M. (2012). Elementary processes governing the evolution of road networks. *Scientific Reports*, *2*. https://doi.org/10.1038/srep00296

Strano, E., Viana, M., da Fontoura Costa, L., Cardillo, A., Porta, S., & Latora, V. (2013). Urban Street Networks, a Comparative Analysis of Ten European Cities. *Environment and Planning B: Planning and Design*, *40*(6), 1071–1086. https://doi.org/10.1068/b38216

Strogatz, S. H. (2014). *Nonlinear Dynamics and Chaos* (2nd ed.). Boulder: Westview Press.

Stumpf, M. P. H., & Porter, M. A. (2012). Critical Truths About Power Laws. *Science*, *335*(6069), 665–666. https://doi.org/10.1126/science.1216142

Suetani, H., Soejima, K., Matsuoka, R., Parlitz, U., & Hata, H. (2012). Manifold Learning Approach for Chaos in the Dripping Faucet. *Physical Review E*, *86*(3). https://doi.org/10.1103/PhysRevE.86.036209

Takens, F. (1981). Detecting Strange Attractors in Turbulence. In D. Rand & L. S. Young (Eds.), *Dynamical Systems and Turbulence* (pp. 366–381). Berlin: Springer-Verlag.

Talen, E. (2011). *City Rules: How Regulations Affect Urban Form* (First American Edition). Island Press.

Talen, E., Menozzi, S., & Schaefer, C. (2015). What is a “Great Neighborhood”? An Analysis of APA’s Top-Rated Places. *Journal of the American Planning Association*, *81*(2), 121–141. https://doi.org/10.1080/01944363.2015.1067573

Theiler, J. (1990). Estimating Fractal Dimension. *Journal of the Optical Society of America A*, *7*(6), 1055–1073.

Tomida, A. G. (2008). Matlab Toolbox and GUI for Analyzing One-Dimensional Chaotic Maps (pp. 321–330). Presented at the International Conference on Computational Sciences and Its Applications ICCSA 2008, Perugia: IEEE. https://doi.org/10.1109/ICCSA.2008.7

Trudeau, R. J. (1994). *Introduction to Graph Theory* (2nd edition). New York: Dover Publications.

Turcotte, D. L. (1999). Self-organized criticality. *Reports on Progress in Physics*, *62*(10), 1377.

Uitermark, J. (2015). Longing for Wikitopia: The study and politics of self-organisation. *Urban Studies*, *52*(13), 2301–2312. https://doi.org/10.1177/0042098015577334

Vitins, B. J., & Axhausen, K. W. (2014). Shape grammars in transport and urban design. In *world symposium on transport and land use research, Delft*. Retrieved from http://www.vitins.ch/papers/WSTLUR\_Vitins\_Axhausen\_2014.pdf

von Hayek, F. (1974, December). *Prize Lecture: The Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 1974*. Nobel Prize Lecture. Retrieved from http://www.nobelprize.org/nobel\_prizes/economic-sciences/laureates/1974/hayek-lecture.html

Walker, B., Holling, C. S., Carpenter, S. R., & Kinzig, A. (2004). Resilience, adaptability and transformability in social–ecological systems. *Ecology and Society*, *9*(2), 5.

Walker, J. (2012). Fractal food. *Artlink*, *32*(1), 33–35.

Wang, F., Antipova, A., & Porta, S. (2011). Street centrality and land use intensity in Baton Rouge, Louisiana. *Journal of Transport Geography*, *19*(2), 285–293. https://doi.org/10.1016/j.jtrangeo.2010.01.004

White, R., & Engelen, G. (1993). Cellular Automata and Fractal Urban Form: a Cellular Modelling Approach to the Evolution of Urban Land-Use Patterns. *Environment and Planning A*, *25*(8), 1175–1199.

Willis, N. (2008, January 23). OpenStreetMap project completes import of United States TIGER data. Retrieved November 4, 2016, from https://www.linux.com/news/openstreetmap-project-completes-import-united-states-tiger-data

Wissen Hayek, U., Efthymiou, D., Farooq, B., von Wirth, T., Teich, M., Neuenschwander, N., & Grêt-Regamey, A. (2015). Quality of urban patterns: Spatially explicit evidence for multiple scales. *Landscape and Urban Planning*, *142*, 47–62. https://doi.org/10.1016/j.landurbplan.2015.05.010

Wolf, A., Swift, J. B., Swinney, H. L., & Vastano, J. A. (1985). Determining Lyapunov Exponents from a Time Series. *Physica D: Nonlinear Phenomena*, *16*(1), 285–317.

Wu, G.-C., & Baleanu, D. (2014). Discrete Fractional Logistic Map and Its Chaos. *Nonlinear Dynamics*, *75*(1–2), 283–287. https://doi.org/10.1007/s11071-013-1065-7

Wu, J., Funk, T. H., Lurmann, F. W., & Winer, A. M. (2005). Improving spatial accuracy of roadway networks and geocoded addresses. *Transactions in GIS*, *9*(4), 585–601.

Xie, F., & Levinson, D. (2007). Measuring the structure of road networks. *Geographical Analysis*, *39*(3), 336–356.

Yeh, A. G.-O., & Li, X. (2001). Measuring and Monitoring of Urban Sprawl in a Rapidly Growing Region Using Entropy. *Photogrammetric Engineering & Remote Sensing*, *67*(1), 83–90.

Zhang, F., Liao, X., & Zhang, G. (2015). Dynamical behavior of a generalized Lorenz system model and its simulation. *Complexity*, n/a-n/a. https://doi.org/10.1002/cplx.21714

Zhong, C., Arisona, S. M., Huang, X., Batty, M., & Schmitt, G. (2014). Detecting the dynamics of urban structure through spatial network analysis. *International Journal of Geographical Information Science*, *28*(11), 2178–2199. https://doi.org/10.1080/13658816.2014.914521

Zielstra, D., & Hochmair, H. (2011). Comparative Study of Pedestrian Accessibility to Transit Stations Using Free and Proprietary Network Data. *Transportation Research Record: Journal of the Transportation Research Board*, *2217*, 145–152. https://doi.org/10.3141/2217-18