Faculty of Science & Technology

Department of Computing & Informatics

Complex Urban Road Networks: Static Structures and Dynamic Processes

Supervisor:
Dr. Wei Koong Chai

Author:
Assemgul Kozhabek

Co-Supervisor:
Prof. Vasilis Katos

This thesis is submitted for the degree of
Doctor of Philosophy

June 2024

Copyright Statement

Dissertation Declaration

This thesis is submitted for the degree of Doctor of Philosophy at Bournemouth University, United Kingdom. I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements. This dissertation contains fewer than 25000 words and has 55 figures and 9 tables.

Dedication

To my son.

Acknowledgements

After finally completing my doctoral dissertation, I reflect on the past four years filled with joy and distress, and I feel grateful to many individuals. Foremost, I express my appreciation to my supervisor, Dr. Wei Koong Chai, for his patience and consistent support and guidance.

I am thankful to the BCP Council and Bournemouth University for their trust in my academic endeavours and for funding my education, as well as to all the faculty and colleagues who have contributed to my journey.

I extend my gratitude to my parents, who have always encouraged me to become the best version of myself and for being role models in loving, caring, and achieving.

Last but not least, I thank my husband and my son for standing by me through all the challenges and successes. Diyar is our sunshine on cloudy days.

Abstract

We first abstract road networks as complex systems and employ tools from network science to study their topological properties. Our multi-scale analysis includes macro-, meso-, and micro-scale perspectives, deriving insights into both common and unexpected patterns in these networks. At the macro-scale, we examine the global properties of these networks. We find correlations between various metrics and capture aspects such as the efficiency and robustness of the network. At the meso-scale, we explore the existence of a sub-structure embedded within the road networks using two main concepts, namely community and core-periphery structures. We found that while these densely populated city road networks show particularly strong community structures (high modularity values that are not typical to other networks), they exhibit a low level of presence of core-periphery structures. This points to the cities being polycentric. Finally, at the micro-scale, we find nodal-level properties of the network. Specifically, we compute the various centrality measures and examine their distributions to capture the prevalent characteristics of these networks. We find different but consistent distributions for the considered centrality measures.

Then we delve into a comparative analysis of the efficacy of various node removal strategies in terms of inducing damage to the reliability of real-world urban road networks. The impact of five strategic (deterministic) based on centrality measures and two random (stochastic) node removal strategies on network robustness is assessed through an iterative node removal process. We assess the robustness of road networks of densely populated cities using three different metrics: size of the largest connected component, global efficiency, and local

efficiency. Our findings suggest that targeted disruptions utilising centrality measures are more effective in disrupting the network than random ones. However, some centrality measures have a strong correlation with each other and thus, requiring combinations of different removal orders to gain more comprehensive insights into the ability of the network to withstand perturbations. We find centrality measures considering shortest paths are more effective in degrading the robustness of the network as a whole while centrality measures that only consider immediate connectivity are better in disrupting the local effectiveness of the network. Interestingly, we also find that removing nodes can counter-intuitively increase the local efficiency of the network.

Lastly, we advocate the use of epidemic theory to model the spreading of traffic congestion in urban cities. Specifically, we use the Susceptible - Infected - Recovered (SIR) model but propose to explicitly consider the road network structure in the model to understand the contagion process of road congestion. This departs from the classical SIR model where homogeneous mixing based on the law of mass action is assumed. For this purpose, we adopt the N-intertwined modeling framework for the SIR model based on continuous-time Markov chain analysis. In our evaluation, we used two real-world traffic datasets collected in California and Los Angeles. We compare our results against both classical and average-degree-based SIR models. Our results show better agreement between the model and actual congestion conditions and shed light on how congestion propagates across a road network. We see the potential application of insights gained from this work on the development of traffic congestion mitigation strategies.

In conclusion, our study comprehensively examines urban road networks by analysing both their static and dynamic aspects.

Keywords: Urban transport networks, network properties, network robustness, largest connected components, global efficiency, local efficiency, congestion spread modeling, epidemics, SIR model, topology, traffic congestion.

Contents

List of Figures

List of Tables

List of Abbreviations

ANN **A**rtificial **N**eural **N**etwork. xv

CNN **C**onvolutional **N**eural **N**etwork. xv

DL **D**eep **L**earning. xv

GCN **G**raph **C**onvolutional **N**eural Network. xv

GISF2E **GIS F**eatures 2 **E**dgelist. xv

GRU **G**ated **R**ecurrent **U**nit. xv

IPNV **I**nfectious **P**ancreatic **N**ecrosis **V**irus. xv

KNN **K**-**n**earest **n**eighbour. xv

LCC **L**argest **C**onnected **C**omponents. xv

LSTM **L**ong-**S**hort **T**erm **M**emory. xv

MAE **M**ean **A**bsolute **E**rror. xv

Max Maximum. xv

Min Minimum. xv

ML **M**achine **L**earning. xv

NFD **N**etwork **F**undamental **D**iagram. xv

ODEs **O**rdinary **D**ifferential **E**quations. xv

PC **P**ath-**c**ore. xv

PeMS **P**erformance **M**easurement **S**ystem. xv

RMSE **R**oot **M**ean **S**quare **E**rror. xv

RNN **R**ecurrent **N**eural **N**etwork. xv

SI **S**usceptible **I**nfected. xv

SIR **S**usceptible **I**nfected **R**ecovered. xv

SIS **S**usceptible **I**nfected **S**usceptible. xv

UK **U**nited **K**ingdom. xv

URNs **U**rban **R**oad **N**etworks. xv

Math Notations

$\bar{d}$	Average degree
β	Propagation rate
$\frac{di(t)}{dt}$	Rate of change of the number of nodes in the *infected* state over time
$\frac{dr(t)}{dt}$	Rate of change of the number of nodes in the *recovered* state over time
$\frac{ds(t)}{dt}$	Rate of change of the number of nodes in the *susceptible* state over time
γ	Recovery rate
κ	Degree diversity
$\lambda_n(t)$	Traffic flow at node n at time t, defined as $\lambda_n(t) = \frac{v_n(t)}{v_n^{max}}$ where $v_n(t)$ is the average speed and v_n^{max} is the speed limit at node n
$\overline{Q_n(t)}$	Effective infinitesimal generator for node n, incorporating the mean-field approximation
ρ	Tuneable threshold value for congestion level. A node is considered congested if $\lambda_i(t) < \rho$
σ_{jk}	Number of shortest paths between node j and node k in network G.
$\sigma_{jk}(v_i)$	Subset of shortest paths between node j and node k that pass through node v_i in the network G, after removing edge (j, k).
τ_G	Transitivity
A	The $N \times N$ adjacency matrix representing graph G

$a_{i,j}$ The adjacency matrix element, indicating whether there is a link between nodes i and j

$a_{m,n}$ Element of the adjacency matrix representing a link between nodes m and n

c_B Betweenness centrality

c_C Closeness centrality

c_D Degree centrality

c_K Katz centrality

c_L Load centrality

CC_G Clustering coefficient

D Network diameter

d_i Degree of node v_i, which is the sum of all connections of node v_i

E Set of links in the graph

$E[q_{1,2;n}]$ Expected transition rate for node n computed using the mean-field approximation

E_i The number of links connecting the neighbors of node v_i

E_{Glob} Global Efficiency

E_{Loc} Local Efficiency

eff_G Efficiency

$G(V, E)$ An undirected graph where V and E are the set of N nodes and L links

H_G Average path length

I Nodes in the *infected* state (congested and can spread congestion to neighboring nodes)

$i(t)$ Number of nodes in the *infected* state at time *t*

$i_n(t)$ Probability of node n being in the *infected* state at time t

L Total number of links in the graph G

m Meshedness coefficient

N Total number of nodes in the graph G

$P(t)$ Epidemic prevalence, defined as the fraction of nodes in the *infected* state at time t

$PC(v_i)$ Path-Core metric of node v_i

Q Modularity

$Q_n(t)$ Infinitesimal generator of the three-state continuous Markov chain for node n

$q_{1,2;n}$ Transition rate from the *susceptible* to the *infected* state for node n

R Nodes in the *recovered* state (previously congested but no longer congested)

$r(t)$ Number of nodes in the *recovered* state at time *t*

r^T The robustness indicator

R_A Random area

$r_n(t)$ Probability of node n being in the *recovered* state at time t

R_P Random point

S Nodes in the *susceptible* state (not congested but can become congested)

$s(t)$ Number of nodes in the *susceptible* state at time *t*

$s_n(t)$ Probability of node n being in the *susceptible* state at time t

S_{LCC} Size of Largest Connected Components

T_G Number of triples at node v_i in graph G

V Set of nodes in the graph

Chapter 1

Introduction

1.1 Background and Motivation

In the past thirty years, the global urban population has risen from 39% to 52% and is projected to reach approximately 66% of the total population by 2050 (Habitat 2022). Cities act as hubs or marketplaces for goods and services, therefore they offer more economic opportunities and have attracted influxes of citizens seeking their fortune.

Figure 1.1 depicts world population projections for urban and rural regions worldwide. Such urban population growth puts strains on the city infrastructures including the transportation system. The seamless functioning of city services relies heavily on an efficient and robust road network. Similar to most real-world networks, urban road networks evolve and grow over time, and their complexity increases from the rise of interconnectedness, and the dynamic processes involved, as vehicles navigate within the networks' configuration and structure (Boeing 2018).

In the realm of urban road networks, *static structures* encompass the unchanging topological properties and spatial design elements that define the network's layout. Assessing the robustness of a road network to disruptions involves exam-

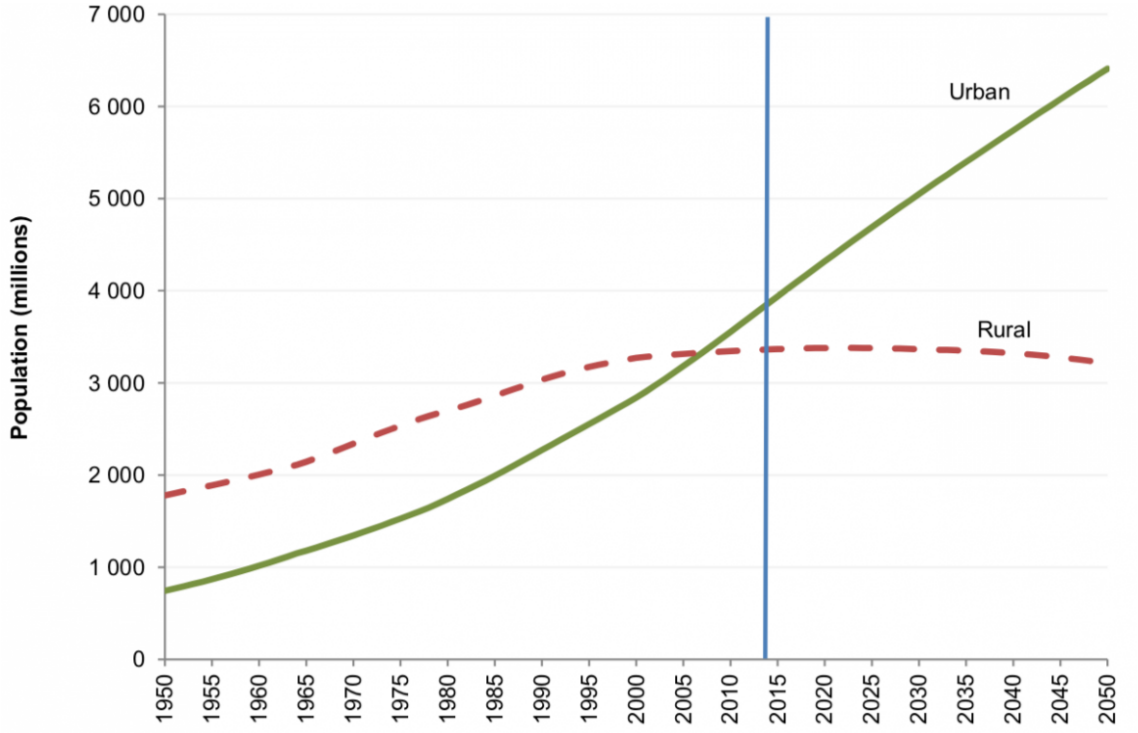


Figure 1.1: Urban population growth trend (UN DESA 2019).

ining how its structure and integrity withstand challenges such as node removal based on predetermined criteria. On the other hand, *dynamic processes* pertain to the time-evolving behaviors, and traffic dynamics, namely congestion propagation and congestion dissemination.

Understanding the static structures of road networks, which encompass persistent topological features outlining their configuration, is essential for a variety of stakeholders, particularly urban planners, transportation engineers, and policymakers. By analysing the topological design of road networks, these stakeholders can make informed choices regarding infrastructure development, traffic control, and sustainable urban development (Badhrudeen et al. 2022; Buhl et al. 2006). Urban planners benefit from an understanding of the network's structure when creating efficient designs that encourage public transportation use, mitigate traffic congestion, and improve overall connectivity within urban areas (Strano et al. 2012; Casali and Heinimann 2019). Transportation engineers utilise topological assessments to enhance traffic flow efficiency and pinpoint critical network nodes for infrastructure enhancements. Policymakers leverage this knowledge to

allocate resources efficiently and prioritise infrastructure initiatives (Tsiotas and Polyzos 2017; Burghardt et al. 2022).

Advancements in understanding the structure and dynamics of urban road networks have been made through the application of various mathematical tools including complex network analysis, graph theory, and epidemic theory to analyse network topology (Boeing 2018) and the corresponding dynamic processes (Kozhabek et al. 2024). Various questions have been investigated using such approaches. These include, for instance, the identification of travel behaviors of city inhabitants (Levinson and El-Geneidy 2009), the evaluation of transportation performance (Levinson 2012), and the comprehension of the urban organisation (Newman 2012; Crucitti et al. 2006). We study the topological properties of real-world urban road networks of the most populated countries.

We map the three prevalent perspectives found in the literature to three scales, namely macro-, meso-, and micro-scale. We employ a collection of metrics and concepts from network science (Barabasi 2013) and draw insights into different properties of these networks and present them in Chapter 3.

The creation of road networks marked a significant technological advancement in human history. Ancient ruins of paved roads have been discovered in the earliest cities of Mesopotamia, dating back to approximately 4000 BC (National Geographic 2024). As human civilizations evolved, the establishment of road systems became integral to the development of urban centres. Expansive road networks facilitated trade between distant cultures and regions known for producing various goods. The Romans, for instance, constructed around 80,000 kilometres of durable highways (Van Tilburg 2007). These road systems play a crucial role in overall infrastructure, enhancing connectivity and accessibility for the transportation of goods and people. Presently, road networks are widespread globally and essential for improving the quality of life by connecting communities and fostering social interactions. It is then not surprising that huge impacts and significant costs are incurred when these road networks do not function well.

The various consequences of inadequate road infrastructure can range from mi-

nor delays due to increased traffic congestion to disruptions or reduced accessibility to essential and emergency services (e.g., ambulance, fire, police), causing both life and economic losses (Ben 2019; Zhang and Cheng 2023). Severe perturbations could also incapacitate some parts of the road networks. Hence, it is important to understand the ability of the network to withstand failures (e.g., for transportation maintenance and planning endeavors). As such, the focus of Chapter 4 is on the robustness assessment of urban road networks. Specifically, we choose ten road networks from densely populated cities and subject them to different perturbations.

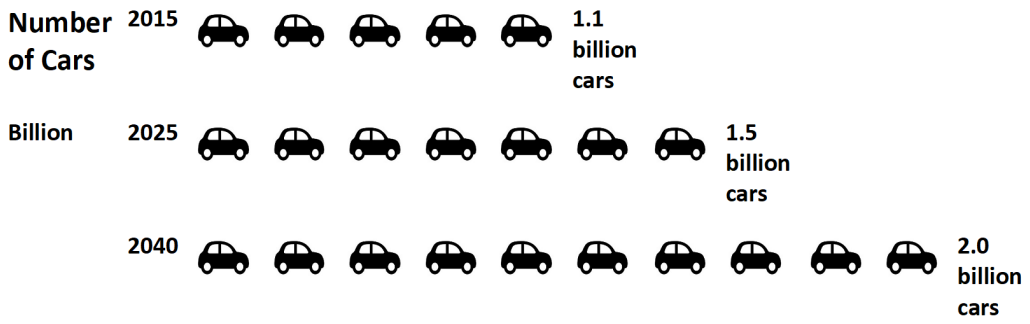


Figure 1.2: Global vehicle ownership growth trend (World Economic Forum 2016).

Despite continuous efforts and technological advancements in recent decades, traffic congestion has been a persistent societal challenge due to dynamic processes such as congestion propagation, which involve time-evolving behaviours in traffic dynamics. The problem is exacerbated by rapid urbanisation (see Figure 1.1) and the growth of motor vehicle ownership to two billion cars by 2040 (see Figure 1.2). Congestion increases journey times, escalates excess fuel consumption, and environmental pollution, and decreases workforce productivity with increasing time loss (Falcocchio and Levinson 2015).

The UK Department for Transport (2022) highlighted that the rise in urban congestion is becoming a substantial problem for cities, leading city transportation

authorities worldwide to introduce various policies and campaigns for congestion mitigation. These measures include motivating people to walk and cycle by introducing congestion charge zones (Transport for London 2023), improving public transport efficiency (Chinese Ministry of Transport 2014), connecting infrastructure with smart traffic signal optimisations (Taipei City Traffic Engineering Office 2019), and utilising dynamic traffic forecasting (Boston Consulting Group 2018).

One crucial component of these efforts involves comprehending the spread of congestion in urban road networks over time and space (Saberi et al. 2020; Nagy and Simon 2021). It offers city municipalities valuable insights into which road segments will be impacted and when. Fundamentally, the question is on how an inherently challenging process like traffic congestion oscillates over time in urban transport systems. From the literature, it is known that congested traffic oscillation itself has two components: formation and propagation (Abdi et al. 2016; Zheng et al. 2023b). The former can be caused by lane-changing activity or any kind of moving bottleneck (Abdi et al. 2016; Li et al. 2020a). The latter is traffic propagation, the subject of Chapter 5, which remains an interesting research topic especially when traffic congestion is a non-linear, non-equilibrium dynamic process that exhibits emergent behaviours (Laval and Leclercq 2010; Chen 2002). Different approaches to model the spread of traffic congestion have been investigated in the literature (see Chapter 2).

To address dynamic processes on urban road networks, specifically, traffic congestion oscillation, we follow the approach based on epidemiology models where traffic congestion evolution is modeled as a disease-spreading process. Basic yet versatile compartmental epidemic models such as the SI (Susceptible - Infected), SIS (Susceptible - Infected - Susceptible), and SIR (Susceptible - Infected - Recovered) models have been applied for simulating the spreading processes in the networks. Wu et al. (2006) proposed to describe traffic congestion spreading with the SIR model whereby a congested road junction or segment is said to be “infected” and conversely, “susceptible” if not congested, and finally, the road junction/segment that becomes uncongested after an episode of congestion can be modeled as a “recovered” state. Furthermore, Saberi et al. (2020) provided em-

pirical validation of traffic congestion propagation and its dissipation as a network epidemic spreading phenomenon and discussed the different possible interpretations of basic epidemic models to represent traffic congestion spreading process. We further extend this line of research in Chapter 5. While previous works have largely taken the basic simplifying assumption of homogeneous mixing of nodes (Anderson and May 1991), in this research, we eschew this assumption and instead explicitly consider the road network topology to model the congestion spreading. Specifically, we adopt an individual-based mean-field approach to model the congestion spreading and apply the model to two real-world traffic datasets.

To sum up, this dissertation delves into the intricate realm of urban road network analysis, focusing on topological insights, structural robustness evaluations, and dynamic traffic congestion modeling. By delving into the structural properties and dynamic behaviors of urban road networks, this dissertation seeks to provide a comprehensive analysis that empowers stakeholders with actionable insights for informed decision-making.

1.2 Aims and Objectives

Influenced by the significance of comprehending the topological characteristics of road networks and their ability to withstand different disruptions, in addition to the growing congestion issue outlined above, this research focuses on three main research objectives below:

OBJ1 Analyse the static structures of complex urban road networks from the macroscopic, mesoscopic, and microscopic points of view. At the macroscopic level, this objective involves analysing the overall layout and overarching features. In the mesoscopic view, the focus shifts to analysing the substructures of URNs. At the microscopic level, OBJ1 involves a detailed analysis of node-level attributes.

OBJ2 Evaluate the robustness of the urban road networks on withstanding (increasing) perturbations. This objective aims to assess how robust the urban

road networks are to various disruptions, such as accidents, road closures, increased traffic volumes, or adverse weather conditions. It addresses the technical challenge of gaining an understanding of how a road network responds to increasing disturbances. We adopt an iterative node removal process for this purpose.

OBJ3 Understanding and modelling of traffic congestion spreading in urban road networks. This objective focuses on studying the dynamics of traffic congestion within urban road networks. The challenge is to develop a model that is capable of accurately capture how traffic congestion evolves over time given a real road network topology.

By addressing these objectives, the research endeavors to shed light on the complex interplay between urban road network structures, traffic dynamics, and urbanisation trends, aiming to provide actionable insights for creating more sustainable, efficient, and robust urban transportation systems.

1.3 Contributions

The main contributions of this dissertation are as follows:

- We present an in-depth analysis of current methodologies for examining urban road network structures, highlighting the importance of structural characteristics and analysing topological properties using network theory. Then we discuss the evaluation of network robustness to various disturbances and explore disruptive scenarios in road networks (cf. Chapter 2).
- We discuss current trends and challenges in modeling traffic congestion, introducing innovative methods like contagion modeling to depict traffic congestion spread at different levels within road networks. We also elaborate on key contagion models (cf. Chapter 2).
- We analyse the topological properties of 29 urban road networks of densely populated cities based on real data. We adopted a network science approach and examined the topological properties of the networks at three

different scales, namely macro-, meso-, and micro-scale (cf. Chapter 3).

- We present insights into specific properties including metrics describing the overall network as a whole, sub-structures within the networks including community and core-periphery patterns, and different distributions of centrality measures that can describe this class of networks (cf. Chapter 3).
- We conduct a robustness assessment based on an iterative node removal process based on both random (stochastic) and targeted (deterministic) disruption strategies (cf. Chapter 4).
- We find that removing nodes based on ranking using path-based centrality measures (e.g., betweenness) offers the greatest impact on the network (cf. Chapter 4).
- We find that removing nodes based on ranking using immediate connectivity (e.g., degree) can have a high impact on local regions but not the network as a whole (cf. Chapter 4).
- We find that due to the similarity/correlation of some centrality measures, there is a need to use a combination of disruption strategies to comprehensively study the robustness of road networks (cf. Chapter 4).
- We model traffic congestion spreading in road networks using epidemic theory, namely a topology-based Susceptible-Infected-Recovered (SIR) model demonstrating the relevance and existence of the spreading phenomena in network traffic congestion dynamics (cf. Chapter 5).
- We conduct a comparison study of traffic congestion dynamics in two real-world urban road networks. The quantitative assessment and comparison were performed on the topology-based SIR model in contrast to classical SIR and average-degree SIR models (cf. Chapter 5).

1.4 Organisation

This thesis is organised into six chapters. The **Chapter 1** introduces the background, motivation and scope, core research works, main contributions, and related publications during the whole PhD research period. The structure of the thesis is summarised in Figure 1.3.

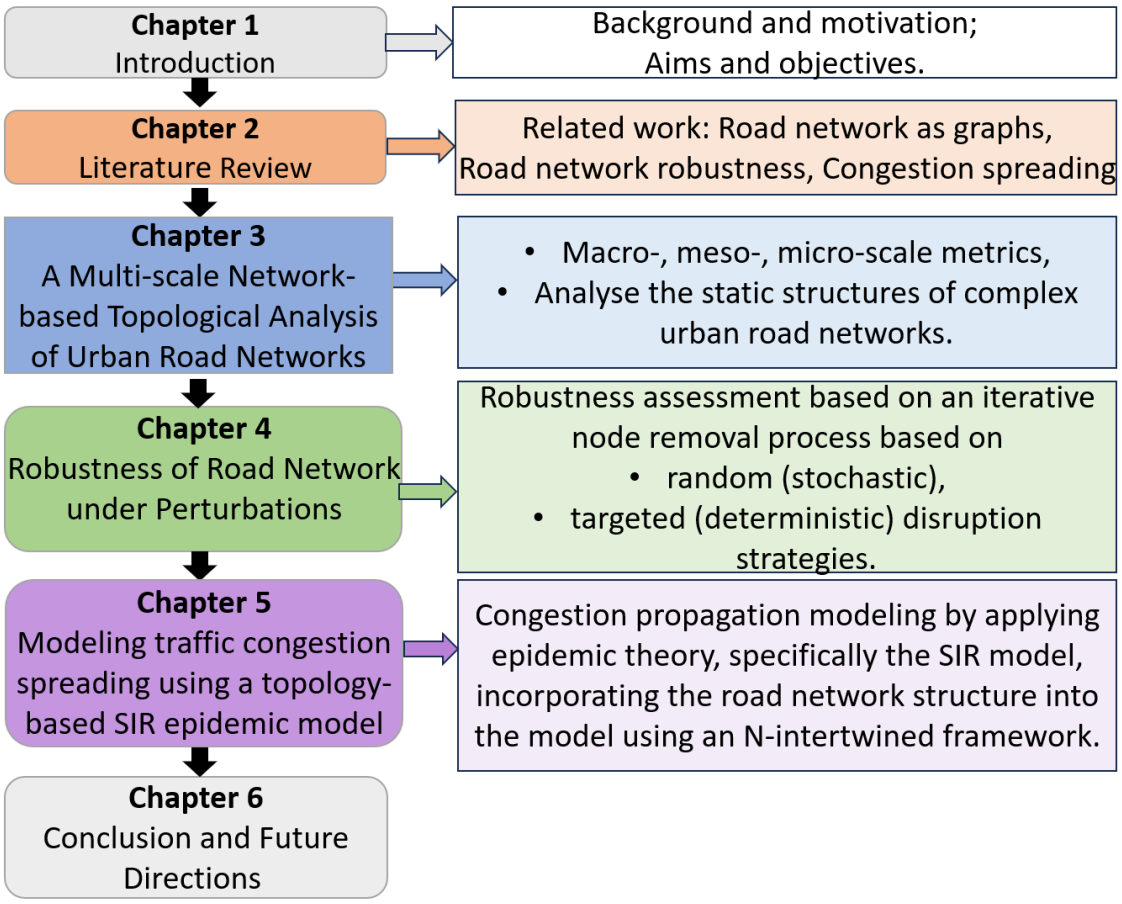


Figure 1.3: Thesis structure.

Chapter 2 first provides some necessary background on network science for this research. It then reviews the relevant state-of-the-art including (1) the present methods for and literature on analysing static topological structures of urban road networks, (2) evaluation of current techniques for assessing the robustness of road networks to different disruptions/perturbations, with emphasis on recognised metrics and (3) the prevailing methods for modeling the spread of traffic congestion.

Chapter 3 examines the urban road infrastructure in highly populated cities, specifically analysing 20 Chinese and 9 Indian road networks. By utilising network science methodologies, the research investigates the topological characteristics of these road networks across macro-, meso-, and micro-scales. The study uncovers overarching features, sub-structures, and node-level attributes, offering valuable insights into the efficiency, resilience, community formations, polycentricity, and centrality metrics of these intricate systems.

Chapter 4 investigates the robustness of urban road networks in densely populated cities by utilising real-world data and conducting a robustness assessment through iterative node removal processes. Various metrics, namely the size of the largest connected component, global efficiency, and local efficiency are monitored during the assessment. The study explores seven node removal strategies, including both stochastic and deterministic approaches based on different centrality measures.

Chapter 5 discusses the problem of traffic congestion in urban areas, linking it to swift urban development and the rise in vehicle possession. By utilising the SIR model derived from epidemic theory to assess the spread of congestion, a unique method is introduced. By incorporating an N-intertwined framework to represent the road network structure, the research aims to enhance comprehension of congestion propagation dynamics. Validation of the model is conducted using actual traffic data, demonstrating promising outcomes in contrast to traditional SIR models.

In the concluding **Chapter 6** of this thesis, the main findings and implications of the research are summarised. Additionally, the chapter discusses potential future research directions that could be pursued based on the outcomes and gaps identified during the study.

1.5 List of Publications and Awards

Papers

1. Assemgul Kozhabek, Wei Koong Chai, and Ge Zheng. 2024. Modeling traffic congestion spreading using topology-based SIR epidemic model. IEEE Access, 12, 35813–35826. doi: 10.1109/ACCESS.2024.3370474
2. Assemgul Kozhabek, and Wei Koong Chai (2024). A Multi-scale Network-based Topological Analysis of Urban Road Networks in Highly Populated Cities. *Environment and Planning B: Urban Analytics and City Science* (Under Revision).
3. Assemgul Kozhabek, and Wei Koong Chai (2024). Robustness of Road Network Structures under Perturbations. *Applied Network Science* (Under Review).

Poster presentation

- Assemgul Kozhabek, Wei Koong Chai, and Ge Zheng. 2024. ”Modeling traffic congestion spreading using topology-based SIR epidemic model”. Bournemouth University SciTech PGR Conference - May 2024.

Awards

* *Scholarship for Events on Complex Systems* (SECS) awardee to attend the Complex Networks 2024 Conference in Istanbul, Turkey (December 10-12, 2024).
* Selected for the *Alan Turing Institute (ATI) Data Study Group* (September 9-13, 2024), a week-long data hackathon with academic and industry ties. Working on a Transport for London project.
* Winner of Poster competition *Green IT - Invest in our Planet* from the British Computer Society (BCS), July 2024.
* *OpenBright* grant awardee for women in computing in December 2023 for

equipment.

* The Winner of the HE Class *GreenIT 2023 Student* competition, organised by the esteemed British Computer Society (BCS) - the chartered institute for IT.

* *OpenBright* grant awardee for women in computing in April 2023 for the "Data Science: Network Science" summer course by Utrecht University, Netherlands.

* Finalist of the TRA (Transport Research Arena) Visions 2024 Young Researchers Competition.

Chapter 2

Literature Review

The literature review begins with an introduction to network science, highlighting its applications in understanding diverse systems. It emphasises the importance of structural characteristics in urban road networks and the utilisation of complex network theory to analyse their topological properties. The chapter also explores the concept of robustness analysis, focusing on the ability of road networks to withstand disruptions during events. The study emphasises understanding urban road networks across macro-, meso-, and micro-scales, particularly in densely populated cities. It covers disruptive scenarios, including random and targeted disturbances, underscoring the significance of comprehending congestion dynamics within urban networks for effective traffic management.

Furthermore, the chapter examines existing literature on modeling methodologies for traffic congestion, highlighting the importance of exploring the spatio-temporal dimensions of congestion propagation. To address the research objectives, innovative methods like contagion modeling are employed to portray traffic congestion spread within a network. Different perspectives on congestion spreading, from vehicle-level to network-level, along with diverse methodologies used for modeling congestion evolution, are explored. The chapter also details contagion models such as the Susceptible-Infected (SI), Susceptible-Infected-Susceptible

(SIS), and Susceptible-Infected-Recovered (SIR) models applied to investigate congestion propagation from a network science stance.

Overall, this chapter lays a strong foundation for examining the dynamic nature of traffic congestion in urban road networks, highlighting the interplay of network structures and the application of contagion models to comprehend congestion spreading.

2.1 Background of Network Science

The field of network science has seen significant expansion in recent years, leveraging the mathematical tools of networks to analyse complex, self-organising systems across social, biological, technological, and economic realms. Pioneering works by Albert and Barabási (2002), Strogatz (2001), and Pastor-Satorras and Vespignani (2004) have led to groundbreaking discoveries in this interdisciplinary domain. Particularly in biological, technological, and infrastructural systems, the study of spatial networks (Barthelemy 2018) has made notable progress, shedding light on their topological, structural, and dynamic characteristics that remain incompletely understood across disciplines like mathematics, physics, biology, and geography.

These networks have proven especially crucial in urban systems (Barthelemy 2016; Pan et al. 2013; Bettencourt and West 2010; Bretagnolle et al. 2006), where investigations have unveiled distinctive city features and notable statistical properties such as scale-invariant patterns observed across diverse urban landscapes (Bettencourt 2013; Goh et al. 2016; Kalapala et al. 2006).

The road network, with its intricate geometry, plays a fundamental role by providing essential connections for residents to navigate through different urban components efficiently. Different street layouts within urban areas offer diverse levels of accessibility, efficiency, and infrastructure utilisation, thereby exerting a significant influence on urban dynamics (Justen et al. 2013; Costa et al. 2010; Wang et al. 2012). As a result, the structural characteristics of road networks have garnered considerable attention in numerous research endeavors, aiming to unravel their

impacts on urban environments and societal functionality (Lämmer et al. 2006; Louf and Barthelemy 2014; Strano et al. 2012).

Furthermore, this exploration delves into the abstraction of road networks as graphs, delving into how principles from graph theory can be effectively employed to dissect and analyse the intricate connectivity and spatial organisation inherent in road systems within urban landscapes. Through this lens, researchers can gain valuable insights into the underlying structure and dynamics of urban road networks, paving the way for informed urban planning and infrastructure development decisions.

2.1.1 Road networks as graphs

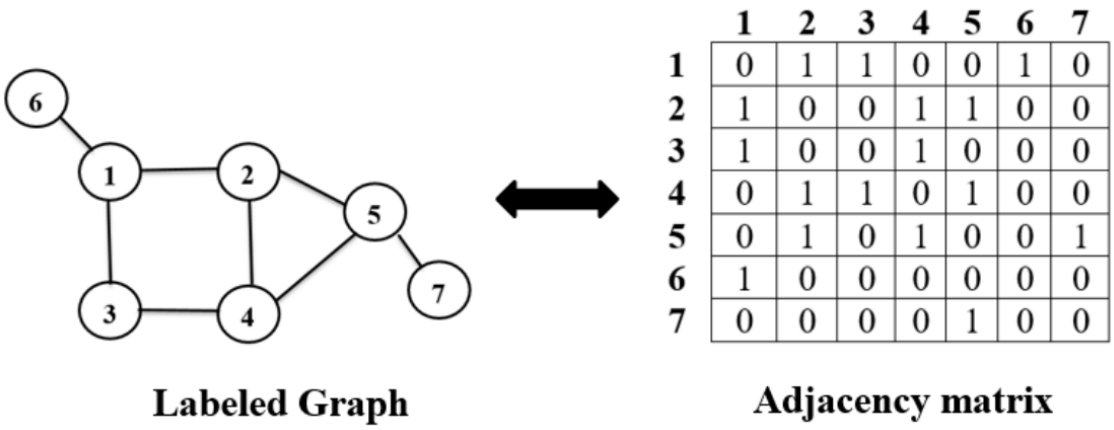
In the context of urban road networks, *static structures* play a crucial role in defining the basic spatial layout and properties of the network that remain relatively constant over short periods of time. These static structures encompass the fixed configurations of road networks, the topological properties of the network, and the inherent characteristics that shape how the network operates.

By treating road networks as graphs in graph theory, we can effectively analyse and interpret the intricate relationships between various components within the network. The concept of utilising graph theory provides a solid framework for studying the relationships between various elements of the network (Blanchard 2009). By abstracting road networks as undirected graphs, denoted as $\mathcal{G}(\mathcal{V}, \mathcal{E})$ with $\mathcal{V} = v_1, ..., v_N$ the set of nodes and $\mathcal{E} = e_1, ..., e_L$ the set of links where $N = |\mathcal{V}|$ and $L = |\mathcal{E}|$.

In our case, the nodes represent the road intersections/junctions while the links represent road segments connecting two nodes. $\mathcal{G}$ can be represented by A, the $N \times N$ symmetric adjacency matrix, with $a_{i,j} = 1$ if there exists a link between nodes v_i and v_j and 0 otherwise. This representation allows us to analyse network properties.

Figure 2.1 illustrates a simple graph with $\mathcal{V}$=1, 2, 3, 4, 5, 6, 7 and $\mathcal{E}$=(1,2), (1,6), (3,4), (2,5), (2,4), (4,5), (5,7) and also its corresponding matrix. By understanding

and studying the static structures of urban road networks through graph representations, we gain insights into how the network is organised, how efficiently it functions, and how robust it is to disturbances or changes.



	1	2	3	4	5	6	7
1	0	1	1	0	0	1	0
2	1	0	0	1	1	0	0
3	1	0	0	1	0	0	0
4	0	1	1	0	1	0	0
5	0	1	0	1	0	0	1
6	1	0	0	0	0	0	0
7	0	0	0	0	1	0	0

Figure 2.1: Adjacency matrix for graph (personal collection 2024).

2.1.2 Topological properties

Aided by the availability of new datasets, there have been various studies on transportation systems including air transportation (e.g., Siozos-Rousoulis et al. 2021; Diop et al. 2021), rail and metro networks (e.g., Wang et al. 2017; Derrible and Kennedy 2010), public transportation networks (e.g., bus network (Li et al. 2020b)) and urban road networks (e.g., Liu et al. 2019; Jiang et al. 2008; Reza et al. 2022; Lämmer et al. 2006; Lee and Jung 2018). These studies have focused on different aspects or questions ranging from understanding topological properties (e.g., computing the efficiency of the network) to finding vulnerabilities or bottlenecks (e.g., identifying key nodes against failure or targeted attacks).

We focus on road networks which prove to be challenging and rich in properties due to the fact that, unlike air, rail, and buses, vehicles do not follow timetables/schedules, and depending on the level of abstractions, the road networks may include various different link types and features (roundabouts vs junctions, highways vs small lanes, traffic lights, etc.). Methodologically, tools and concepts from network science (Barabasi 2013), with roots from graph theory, have widely been used in the examination and understanding of these complex net-

works (Newman 2012). Figure 2.2 represents a sample road network and the graph that is built considering the selected road network.

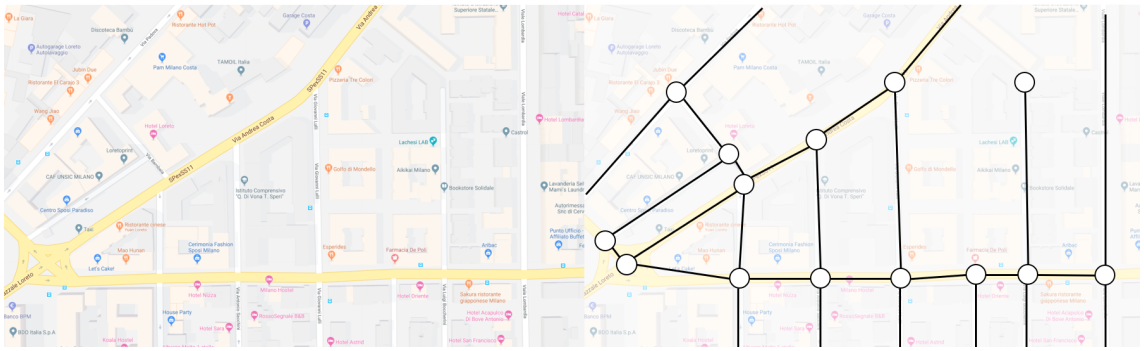


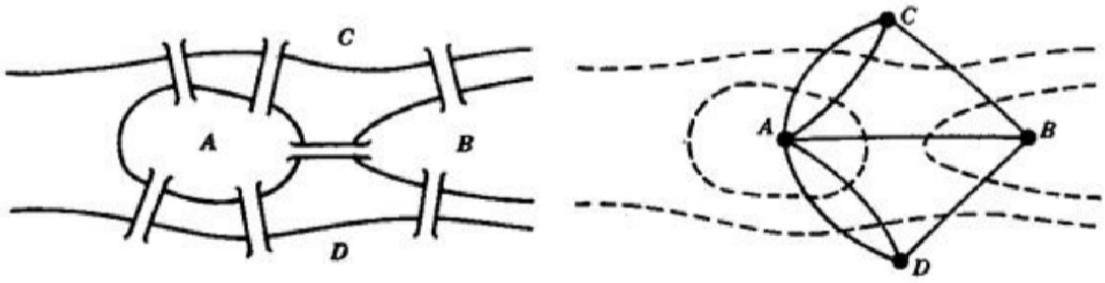
Figure 2.2: Actual road network from Google Maps (2024) (left side) and its representation as undirected graph (right side).

In fact, the well-known Königsberg's bridges problem, solved by Euler in 1735 (Hopkins and Wilson 2004), can be considered as the earliest and simplified transportation optimisation problem (see Figure 2.3). This problem not only served as a basis for the development of graph theory but also introduced the initial concepts of network topology into transportation research (Barabasi 2013). Since then, there are various metrics and approaches developed. The layout of road networks influences efficiency, accessibility, and infrastructure utilisation, contributing to the overall functionality of urban spaces. Understanding the structural aspects of street networks has been a focal point of research efforts, shedding light on their topological characteristics and performance across different urban contexts. This sets the stage for a deeper dive into the specific analysis of urban road networks in the subsequent subsection.

2.1.3 Structural attributes of urban road networks

Urban road networks are intricate systems with a high number of nodes and edges, influenced by factors like node connectivity, importance, and drivers from diverse socio-economic backgrounds. Studying the network's topology is crucial to comprehend its structure and identify key features.

One natural way to study road networks is to view the entire topology as one whole system and compute some overall properties (often taking the mean value



(a) Königsberg bridge (b) Euler's graphical representation

Figure 2.3: In the Prussian city of Königsberg, four regions are known as A, B, C, and D, linked by seven bridges (Figure 2.3a). The challenge involves creating a route that traverses each bridge exactly once. Euler approached the problem by creating a graphical representation (Figure 2.3b) in which each node corresponds to a region and each edge represents a bridge. He demonstrated that it is not possible to complete such a walk (Hopkins and Wilson 2004).

of some indicator). Basic metrics such as network diameter/radius, link density, average degree, average path length, and global clustering coefficient are commonly used. The significance of these metrics has been explored, for instance, in Buhl et al. (2006), Jiang and Claramunt (2004), and Barthélemy (2011). Average node degree is one of the most used metrics (not limited to road networks). It offers a first insight into the level of connectivity in the network. However, road networks are spatial networks, and thus, the physical road infrastructure is often constrained by the actual geographical space and terrain of the geographical area. Studies such as Boeing (2017) and Barthelemy (2018) have all noted that due to spatial restrictions, often the degree is also constrained. On the other hand, studies such as Novak and Sullivan (2014) and Sharifi (2019) highlighted the importance of path-based metrics such as average path lengths which may also be influenced by physical space as well as dynamic events/processes (e.g., congestion, disruptions or road closures). While it is valuable to analyse the network as a whole system and compute overall properties, focusing solely on these macroscopic metrics may overlook crucial insights at smaller scales.

Sub-structures

As opposed to the above-mentioned body of work focusing on the average behavior of the network as one whole system, other studies have focused on the sub-structure of the network. The detection of community structures in a network was first introduced by Newman (2004) where *modularity* is proposed to measure the presence of communities in a network. This has subsequently been applied to road networks (e.g., Song and Wang 2010; Duan and Lu 2013).

Specific to urban cities, Tsiotas and Polyzos (2017) found that modularity has a correlation with regional urbanisation and exploited it as an indicator of the presence of either polycentrism or monocentrism in regional structures. Modularity is also used to find nodes with the closest relations in a road network (Tang et al. 2013). On the other hand, Sun et al. (2016) proposed a spectral method based on modularity to obtain traffic communities in a network focusing on taxi trips which is used to understand the effects of various traffic needs on the spatial distribution of communities. A heuristic approach was employed by Yildirimoglu and Kim (2018) to maximize modularity based on the interactions between different traffic types.

The core-periphery structure is another sub-structure often found in networks and could describe the existence of highly mesh nodes in the centre of a city. Works such as Cucuringu et al. (2016) and Lee et al. (2014) extend the application of this concept to road networks, developing the idea of transport-based coreness instead of the original concept purely based on connectivity as in Borgatti and Everett (2000). Examining the sub-structure of the network offers insights into the intermediate structures and connections within the system.

Nodal-level properties

Meanwhile, there is also another body of work focusing on nodal-level properties of a network (i.e., addressing the importance or role of individual nodes). These works commonly use a combination of different centrality measures to answer their research questions. Degree, betweenness, and closeness centralities ap-

pear to be the most used measures (e.g., Crucitti et al. 2006; Liu et al. 2019; Reza et al. 2022; Lin and Ban 2017; Porta et al. 2006; Shang et al. 2020b) in this category of studies. In Crucitti et al. (2006), the authors conducted an analysis of centrality measures and their spatial distributions across urban networks in different cities around the world. Centrality is also used to examine growth patterns as a result of urbanisation. For instance, the evolution of Paris from 1789 to 2010 was investigated, highlighting a reorganisation in the spatial distribution of centrality (Barthelemy et al. 2013). In Strano et al. (2012), authors identified two urbanisation processes, densification and exploration, in a northern area of Milan, using differences in betweenness centrality values at each edge to quantify these processes.

The relationship between centrality measures in road networks and land use was also explored (e.g., in Wang et al. 2011; Zeng et al. 2020). These works found that closeness centrality showed the highest correlation with land use densities, while betweenness centrality exhibited a weaker correlation. In Jiang et al. (2008), authors investigated the correlation between traffic flow and centrality measures in self-organised road network systems. In the city of Bologna, Porta et al. (2009) studied the connection between the centrality of the street interconnections and the densities of commercial and service activities. They discovered a high correlation between these activities and the global betweenness of the street network, as well as a slightly lower correlation with the global closeness centrality.

From the above, we can already see the literature has adopted different perspectives when studying road networks. Employing aggregated network-wide metrics offers insights into the network property as a whole while nodal-level metrics indicate the importance of nodes in the network. As such, the different perspectives have their strengths. In this research, we thus advocate for a multi-scale analysis approach to understand the full range of road network characteristics.

The preceding subsection delves into the representation of road networks as graphs, where key topological properties are examined to understand their structural characteristics and performance. Transitioning from the conceptual framework of road networks as graphs, the focus now shifts towards conducting ro-

bustness analysis to evaluate the network's resilience under various disturbances and disruptions. This exploration will illuminate the network's ability to maintain its functionality and connectivity amidst challenges, offering valuable insights into enhancing urban road network designs and infrastructure management strategies. This analysis of robustness paves the way for a detailed examination of specific aspects related to urban road networks in the subsequent subsection.

2.2 Robustness Analysis

Network robustness has been a staple topic of study in network science over the past decades and has been applied to various types of complex real-world networks. While the concept can be easily intuited, a formal standardised definition proved to be more elusive and highly dependent on the type of network or application of interest (Smith et al. 2011).

Additionally, scholars have encountered challenges in reaching a consensus on the definition of robustness within the realm of networked systems. Various terms such as reliability, survivability, safety, and resilience have similar meanings to robustness, further complicating consensus-building among researchers. In Immers et al. (2004), robustness is described as "the extent to which a system can operate in line with its design specifications even during significant disruptions". Meanwhile, Boccaletti et al. (2006) defined robustness as a network's capacity to function despite damage to some of its components. Robustness in Schillo et al. (2001) appears as the capability to uphold "safety responsibilities", while also linking it broadly to system performance. In the context of road networks and for our study, we will adapt the definition by Van Mieghem et al. (2010), where *robustness is the network's ability to withstand perturbations under node disruption*.

Several studies have explored the evaluation of network robustness in various disaster scenarios, aiming to enhance the robustness of urban road networks. In particular, researchers have introduced different methodologies and indices to analyse the robustness of urban road networks under different conditions, with a focus on minimising the isolation of districts during disruptive events. Sakak-

ibara et al. (2004) introduced a topological index grounded on the concentration of links in a network to evaluate robustness in disaster scenarios, to reduce the network breaking into multiple components during these events. Similarly, Scott et al. (2006) defines and applies the network robustness index, for evaluating the critical importance of a link in three hypothetical road networks. They measure robustness by calculating the change in travel time when a link is completely removed. Zhou et al. (2017) have introduced a framework for evaluating the robustness of urban road networks, consisting of two layers. The framework includes variations for measuring robustness against random failures and intentional disruptions and is validated using a real-world urban road network in Hong Kong.

The exploration of network robustness in various contexts has presented a complex landscape shaped by diverse perspectives and interpretations. Definitions of robustness have evolved to encompass notions of reliability, survivability, safety, and resilience, each offering unique insights into the capacity of systems to withstand disruptions. Scholars have grappled with aligning on a singular definition, with nuances arising from the specific characteristics of the network under examination. For our study on urban road networks, we adopt a definition of robustness that emphasises the network's ability to endure perturbations in the face of node disruptions. With a focus on enhancing the robustness of urban road networks in the face of disasters, researchers have introduced innovative methodologies and indices to evaluate network robustness under various disruptive scenarios.

Disruptive scenarios

Disruptive events have the potential to influence both the availability of road networks and the preferences of users. Damage to road network structure can impact the supply side of the transport system, leading to disruptions in traffic flow due to road closures and reduced speed and capacity. On the other hand, the demand side, which includes user behaviors such as travel patterns and average speeds, is also affected. This can result in trip cancellations due to affected destinations or the need to find alternative routes when faced with congestion on typical roads. Disruptive scenarios were studied through random or targeted

disturbances.

To illustrate, Mayor of London (2022) reported that congestion costs the London economy £5.1bn a year (equivalent to £1,211 per driver). On average, drivers in India, one of the most populous countries in the world, waste 135 hours per year in traffic congestion per year, costing USD$22bn involving fuel waste, high air pollution, and productivity loss (UN ESCAP 2021). The collapsed Zijin Bridge in Heyuan City in June 2019 led to significant direct and indirect economic losses, limited mobility, and incalculable ecological imbalances (Tan et al. 2020).

Random disturbances

The random strategy removes random network nodes, whereas the targeted strategy removes network nodes based on a classification of their centrality values (Scardoni and Laudanna 2012; Casali and Heinimann 2020; Iyer et al. 2013; Bellingeri et al. 2020; Ma et al. 2020; Kumar and Singh 2020). In Buhl et al. 2006, the authors investigated the robustness of the road network by progressively removing nodes in a random way and by decreasing the order of node degree in 41 urban settlements.

Random disturbances can manifest in various forms, such as:

- Natural disasters - earthquakes, hurricanes, tsunamis and floods can damage road infrastructure, leading to closures and rerouting of traffic (e.g., due to extreme heavy rainfall in Zhengzhou city, China, road links were disrupted (Stanway 2021); the monsoon floods caused the disappearance of numerous bridges and the devastation of over 10,000 km of roads in Kerala, India (Parthasarathy et al. 2021)).
- Accidents and activities - vehicle collisions on major roadways can result in temporary closures and traffic congestion, impacting travel times and network performance human activities (e.g., day-long farmers strike in Delhi caused massive traffic congestion on main roads (The Economic Times 2024); a catastrophic traffic jam in Jakarta, Indonesia, which originated from a traffic accident, persisted for 35 hours and resulted in grave outcomes in

July 2016 (Webb 2016). Zheng et al. (2020) found that the level of traffic congestion is predominantly influenced by the characteristics of traffic incidents, the vehicles involved, and the timing of the events;

- Unexpected maintenance issues - sudden infrastructure failures or the need for urgent repairs can disrupt normal traffic flow and require immediate attention to restore functionality (e.g., a recent bridge collapse in Baltimore has led to severe consequences, greatly impacting commuters preparing for significantly longer travel times lasting three to four times as long as usual (Craig et al. 2024); in September 2022, a substantial sinkhole opened up on the main road in Tel Aviv, causing extensive traffic disturbances throughout the city (The Times Israel 2022)).

These disturbances can lead to road closures, detours, and changes in travel times, affecting the overall efficiency and reliability of the road network.

Targeted disturbances

Targeted road network disturbances involve deliberate actions to disrupt specific areas of the road network based on predefined criteria. These disturbances are strategically planned to impact the transportation system in a specific manner, often aiming to create specific effects or test the robustness of the network under controlled conditions (i.e., scheduled maintenance such as patching potholes on city roads, resurfacing, and painting lane markings). Generally, targeted disturbances are frequently characterised by a reduced likelihood of happening yet can result in catastrophic consequences.

A study by Albert et al. (2000) utilises metrics like variations in diameter, the size of the largest cluster, and the average size of isolated clusters to evaluate how networks respond to intentional or random disruptions. Another study in Duan and Lu (2014) analysing six-city road networks globally found that the diversity in betweenness centrality distribution of the network plays a key role in its robustness, more than geographical. While studied road networks in this study exhibit similar performances under disruptions due to shared topological structures, but differ by granularity because city road networks modeled at different granularities

have different topological structures. These studies have contributed to the development of a detailed analysis of how road network robustness properties may change under a broad range of alterations. Our work here follows the same line of approach but focuses on road networks in densely populated cities which have some specific but prevalent properties such as lack of degree diversity due to the physical spatial constraint in the construction of road infrastructure.

Whether through random or targeted disturbances, the robustness of road networks are put to the test. These disruptions can lead to road closures, traffic rerouting, and changes in travel patterns, ultimately influencing the network's efficiency. The next section delves into dynamic processes in the URNs focusing on congestion propagation.

2.3 Congestion Spreading

Dynamic processes pertain to the evolving and time-dependent phenomena within urban road networks, involving the movement of vehicles, traffic flow patterns, congestion dynamics, and the propagation of traffic disruptions or incidents. These processes capture the temporal behaviors and interactions that occur within the road network. Here and further, we focus on the congestion propagation.

Traffic congestion is a significant issue in large urban areas, as studies have shown that simply building more roads does not alleviate the problem (Glaeser 2011; Bansal and Graham 2023). Transportation engineers and planners rely on traffic models to help them design and implement effective measures to reduce congestion. However, accurately modeling congestion and implementing optimal traffic control strategies in urban networks poses a persistent challenge. Congestion spreads spatially and temporally, requiring computationally intensive simulations to understand patterns and design efficient control strategies. Recognising the spatio-temporal nature of congestion propagation is crucial for traffic management. Factors such as travel demand, infrastructure, and control strategies influence congestion, with high spatial correlation often leading to congestion in adjacent roadways (Li et al. 2020a).

The literature on road congestion formation and spreading has a long history and the topic has been investigated from different perspectives. We broadly segregate them into three levels, namely vehicle-level, link-level, and network-level. Understanding congestion at these levels can provide insights into the factors contributing to congestion and help in devising targeted solutions to manage traffic flow efficiently. At the vehicle-level, factors like acceleration and deceleration patterns impact traffic flow dynamics. At the link-level, characteristics of individual road segments influence how congestion propagates through the network. Finally, at the network-level, the overall network structure and interconnectedness play a critical role in congestion spreading. By examining congestion from these multiple levels, researchers and practitioners can develop comprehensive strategies to address congestion challenges in urban areas.

2.3.1 Vehicle-level congestion spreading

Work adopting the vehicle-level perspective incorporates individual vehicle's movement and its interactions with other vehicles into consideration. Ahmed et al. (2021) identified three approaches to model these interactions, including:

(1) Car following models (e.g., Pipes 1953; Brackstone and McDonald 1999; Saifuzzaman and Zheng 2014) where the models take into account how drivers behave when they are driving behind another vehicle in proximity,

(2) Lane changing models (e.g., Rahman et al. 2013) where the models endeavour to capture the idea of replicating the driver's decision to change lanes influenced by diverse factors,

(3) Gap acceptance models (e.g., Akçelik 2007) where the models attempt to mimic human behaviour and take into account the reasons for driving decisions.

However, these models focus more on the question of congestion formation but are not able to provide further insights into how congestion spreads in a road network.

2.3.2 Link-level congestion spreading

The work adopting the link-level view focuses on individual road segments instead and studies traffic as a continuous and aggregated stream. One direction adopting this view employs queueing theory in which the road segment is modeled as a queue with the vehicles traveling across it as the customers (Vickrey (1969), Arnott (2013), and Zhu et al. (2022)). Alternatively, Lighthill and Whitham (1955), Richards (1956), and Newell (1993) advocated kinematic wave theories to describe the vehicles' behaviour across a road segment in which traffic is modeled as shockwaves in a two-dimensional time-space diagram either as forward or backwards moving.

Such models are often applied to understand traffic on a single freeway and can be extended to small road networks. While they offer insights into the traffic evolution within a road segment, they are unable to capture the complex congestion evolution or behaviour of large-scale urban road networks, such as those found in major cities, where the road network topology offers rich possibilities for traveling routes as well as opportunities of disruptions.

2.3.3 Network-level congestion spreading

To understand how traffic congestion evolves over time in large-scale road networks, a network-level perspective is adopted. The structure of road networks is known to affect the traffic flow at the city scale (Kurant and Thiran 2006; Olmos and Muñoz 2017; Verbavatz and Barthelemy 2020). Levinson (2012) underlined the importance of understanding and monitoring urban street network structures for planning and evaluation of the effectiveness of investments and land use changes within cities. Considering the network-level traffic congestion spread is more challenging as it evolves in multiple directions over space and time.

Some researchers attempted to model this by abstracting it as a physical system (Piccoli and Garavello 2006; Barabási 2016). For instance, Payne (1971) and Whitham (1974) developed a traffic flow model using fluid dynamics law where the

flow of vehicles is represented as the flow of liquid in a pipe. Helbing (1996) developed a gas-kinetic equation-based macroscopic traffic model, while Vandaele et al. (2000) proposed the use of speed-flow and speed-density diagrams for this purpose. Mahmassani et al. (2013) investigate the application of the network fundamental diagram (NFD) as a means to observe, comprehend, and model the loading and unloading processes of urban traffic. Furthermore, the authors of Zeng et al. (2019), Li et al. (2015), and Anbaroglu et al. (2014) explore congestion propagation through the lens of network theory, assessing the severity of congestion through clustering methods. However, these works do not address the question of understanding the dynamics of congestion propagation which we look into in this research.

Meanwhile, recent advancements in artificial intelligence have given rise to new machine learning (ML) and deep learning (DL) models developed for traffic prediction. We refer readers to Zheng et al. (2023a) and the references therein for an overview of the evolution of the research along this direction from simple ML approaches (e.g., K-nearest neighbour (KNN) algorithm (Ko et al. 2016), Artificial Neural Network (ANN) (Huang 2014; Zhu et al. 2018)) to DL models (e.g., Recurrent Neural Networks (RNN) and its variants (Long-Short Term Memory (LSTM) and Gated Recurrent Unit (GRU) (Fu et al. 2016), Convolutional Neural Network (CNN)(Toncharoen and Piantanakulchai 2018), and Graph Convolutional Neural Network (GCN) (Diehl et al. 2019)). Here, rather than understanding the congestion spreading, the focus is on predicting the state of the road traffic in the next prediction horizon (e.g., could be set to 5-min or 10-min in the future) based on historical data. The models, especially the recent ones, are usually complex, entail high computational costs, and are dependent on the size and quality of the dataset.

Finally, traffic congestion spreading has also been investigated as a contagion process (e.g., Zheng et al. 2007, Chen et al. 2022, Li et al. 2017b, Zhang et al. 2022, Zheng et al. 2023b). In this approach, the road network is abstracted and represented as a graph consisting of a collection of nodes (road junctions/intersections in our case) and links (road segments). Traffic congestion evolution can

then be replicated as a spreading process on the graph via tools such as epidemic theory which originated from biological studies on contagious disease spreading and has found applications in various domains (e.g., in communication networks (Meloni et al. 2009), in (online) social network (H. et al. 2014), in hub protein and brain structure (Iturria-Medina et al. 2014)).

The analogy in our case is quite intuitive where a congested road segment can "infect" its adjacent road due to traffic queues extending and overflowing to the next junctions. Colizza et al. (2007) associated the epidemic spreading with the traffic flow based on the metapopulation model. Wu et al. (2004), on the other hand, described traffic congestion spread with the SIR epidemic model and conducted simulation studies. They highlighted the importance of considering the topological properties of the road network in affecting the behaviour of the traffic system.

Furthermore, there is already evidence in the transportation domain that congestion propagation patterns show similarity with the virus contagion process in the real world (Jiang and Claramunt 2004, Brockmann and Helbing 2013, and Sahneh et al. 2013). Specifically, Saberi et al. (2020) used empirical data from Google for different cities and showed that congestion spreading behaves akin to an epidemic process. Following the above, the literature has highlighted the potential of the epidemic model in replicating congestion spreading, and at the same time, the model should take into account the road network while not incurring expensive computational costs or requiring a large volume of data.

2.3.4 Contagion Models

The rapid spread of any phenomenon across a network of connected entities is commonly referred to as contagion processes. Typically, contagion is used to analyse and predict the outbreak and transmission of infectious diseases or viruses, as well as to devise strategies to prevent or control the spread within the network. Traditional epidemiological models assume random interactions within the population network. However, each member of the network has a limited number of contacts through which an infection can spread (Keeling and Eames

2005). Therefore, employing a network-theoretic approach to model infectious disease transmission has been shown to be more precise.

The mechanisms of spreading contagious infections or viruses within a contact network are relatively well understood (Brockmann and Helbing 2013; Shulgin et al. 1998). Various analytical and numerical frameworks have been proposed in the literature to study the propagation of different spreading phenomena, such as computer viruses on the internet or the dissemination of opinions on social media. The framework of contagion-based modeling is built upon the central idea that each element in the network is segregated based on its infection status in relation to the contagion process.

Three standard contagion models (SI, SIS, and SIR) have been developed and are elaborated on in the subsequent sections, with their applicability depending on the nature of the contagion being studied. These models are also utilised to assess the effectiveness of links in transmitting information within a vast social network (Zhang et al. 2018). Understanding the importance of these links in information dissemination, where each link varies in its diffusion capacity can help identify influential links that can either hinder or slow down the spread of information. This process of identifying key links for diffusion control is termed "Transmission Centrality" and is based on the Susceptible-Infected (SI) model.

Susceptible-Infected (SI) model

This model operates on the basic premise that a network node may either be infected or have a likelihood of contacting an infected node and becoming infected itself. In this model, it is asserted that once a person or node is infected, there is no possibility of treatment or recovery, and the node will stay infected indefinitely. In a scenario where the network exhibits homogeneous mixing (where every node has an equal number of connections), all nodes are vulnerable to infection at the onset $t = 0$. The commencement of infection occurs with the introduction of a single infected individual, and the progression of infections can be predicted at any given time by employing the methodology outlined in Barabasi (2013).

$$\frac{di}{dt} = \beta i(1-i) \tag{2.1}$$

In this model (i.e., Eq. 2.1), i represents the proportion of infected nodes in the network at time t, while β denotes the propagation rate and k signifies the nodes' degrees. The infection spreads from an infected individual to a susceptible one with a rate of β. The model posits that recovery from infection is not possible, implying that the infection is enduring. For instance, this model was effectively applied in Scottish salmon farms to simulate the transmission of the Infectious Pancreatic Necrosis Virus (IPNV) (Murray 2006).

Susceptible-Infected-Susceptible (SIS) model

Similar to the SI model, SIS also features only two infection states for a node at any given time t. However, it focuses on illnesses or contagions that have a recovery or cure phase, allowing an infected node to revert to a susceptible state after a period. By maintaining homogeneous mixing and the initial conditions from Eq. 2.1, the infection rate can be determined, as outlined in reference Barabasi (2013):

$$\frac{di}{dt} = \beta i(1-i) - \gamma i \tag{2.2}$$

Introducing the additional parameter γ as the recovery rate of nodes, the term γi in Eq. 2.2 illustrates how quickly the population recovers post-contagion exposure. With the possibility of achieving cures, the system reaches an endemic state over time i, where the percentage of infected nodes $i(t)$ remains constant. It is important to highlight that a higher recovery rate γ leads to a rapid decline in the number of infected nodes, ultimately eradicating the contagion. This model has been widely utilised in the past to depict the transmission of diseases such as diphtheria, chlamydia, or gonorrhoea (Keeling and Eames 2005; Hethcote and Yorke 1984).

Susceptible-Infected-Recovered (SIR) model

In SIR modeling, nodes are primarily classified into the following stages:

- Susceptible (S): Nodes with a vulnerability to encountering the contagion.
- Infectious (I): Nodes that have been exposed to the contagion.
- Recovered (R): Nodes that have either regained full immunity after being infectious or have been eliminated from the network.

This model introduces a new infection state for nodes that have been cured or recovered, transitioning them to a recovered status where they are no longer susceptible to infection. Unlike the SIS model, the SIR model proposes that recovered nodes or individuals move into a recovery state, signifying the development of immunity rather than remaining vulnerable to reinfection. This assumption accommodates contagions or diseases that manifest only once within a specific time frame or afflictions that lead to fatality without available cure, such as measles, whooping cough, or smallpox (Keeling and Eames 2005; Barabasi 2013). With consistent network assumptions, the following derivation applies:

$$\frac{di}{dt} = \beta i(1 - r - i) - \gamma i \tag{2.3}$$

where r represents the fraction of recovered nodes in the network at any time t. Similarly, we can obtain the change in susceptible and recovery nodes:

$$\frac{ds}{dt} = -\beta i(1 - r - i) \tag{2.4}$$

$$\frac{dr}{dt} = \gamma i \tag{2.5}$$

These differential equations outline the temporal progression of the proportions of individuals in susceptible (s), infected (i), and removed/recovered (r) states. Similar to the SIS model, the status of infected nodes within the network is influenced by the transmission rate (β) and recovery rate (γ), along with a fraction

of recovered connections. Hence, the rates of susceptibility (Eq. 2.4 & 2.5) and recovery (Eq. 2.5) directly impact the rate of infection (Eq. 2.3). The total sum of individuals in these states at any given time (t) remains constant, as expressed by the equation $i(t) + s(t) + r(t) = 1$.

The SIR model and its variations have been effectively utilised in studying infectious disease spread and generating disease control strategies. However, the application of this model extends beyond epidemics to include contagion modeling in social and digital networks, such as examining the dissemination of opinions on platforms like Facebook (Rui et al. 2018) and Twitter (Abdullah and Wu 2011), or the spread of computer viruses within Internet networks (Nika et al. 2015; Opuszko and Ruhland 2013).

To sum up, the section navigates through the mathematical formulations and applications of these models in simulating infectious disease transmission, emphasising their relevance in not only epidemiology but also in studying congestion dissemination and network behaviors in various contexts.

2.4 Chapter Remarks

The literature review chapter has provided an extensive overview of current methodologies for analysing urban road network structures, evaluating network robustness to traffic disturbances, and modeling traffic congestion propagation. The chapter starts by introducing network science and its applications in understanding urban road networks. Then it highlights the significance of the structural characteristics of urban road networks and discusses the use of complex network theory to analyse their topological properties. It further explores the concept of robustness analysis, examining the ability of road networks to withstand perturbations under disruptive events.

Literature has adopted various different perspectives when studying road networks. Employing aggregated network-wide metrics offers insights into the network property as a whole while nodal-level metrics give indications on the importance of nodes in the network. As such, the different perspectives have their

strengths. We thus see a gap in which most relevant studies are still studying road networks focusing on one particular perspective or even specific metric. This research is then motivated by the opportunity and potential to exploit the advantages offered by different perspectives in studying road networks. Building on this motivation, urban road networks need to be understood at different levels of abstraction to appreciate the hidden topological properties and their impacts. Hence, in this study, we conduct the analysis across three scales—macro-, meso-, and micro-scales—targeting specifically densely populated cities.

The literature review also discusses current trends and challenges in modeling traffic congestion. While traditional modeling techniques effectively capture overall congestion patterns in a network, there is a need for further investigation into the spatio-temporal aspects of congestion propagation and dissipation. To meet the research objectives, this dissertation employs innovative methods like contagion modeling to depict the spread of traffic congestion within a network. It discussed the different perspectives on congestion spreading, ranging from the vehicle-level to the network-level, and the various methodologies employed to model congestion evolution. Additionally, the chapter discussed contagion models, including the Susceptible-Infected (SI), Susceptible-Infected-Susceptible (SIS), and Susceptible-Infected-Recovered (SIR) models, to investigate congestion propagation from a network science perspective.

Overall, the literature review chapter sets the foundation for exploring the dynamic nature of traffic congestion in urban road networks, the interplay of network structures, and the application of contagion models to understand congestion spreading. This comprehensive analysis contributes valuable insights to the field of transportation engineering and urban planning, paving the way for further research in optimising traffic management strategies and enhancing the resilience of urban road networks.

Chapter 3

A Multi-scale Network-based Topological Analysis of Urban Road Networks

This chapter delves into the analysis of urban road infrastructure within densely populated cities, specifically focusing on 20 Chinese road networks and 9 Indian road networks. Utilising methodologies from network science, the study embarks on exploring the topological characteristics of these road networks across macroscopic, mesoscopic, and microscopic scales. By examining global properties, sub-structures, and nodal-level attributes, valuable insights are extracted about the efficiency, robustness, community structures, polycentricity, and centrality measures of these intricate systems. The research outcomes not only offer a deep understanding of the underlying dynamics shaping these urban road networks, but also provide a framework for informed decision-making and strategic city development planning.

The illustration in Figure 3.1 delineated metrics at macro-, meso-, and micro-scales. By consolidating metrics from these distinct scales, a more comprehensive understanding of the architecture can be achieved, enabling a detailed as-

sessment and analysis.

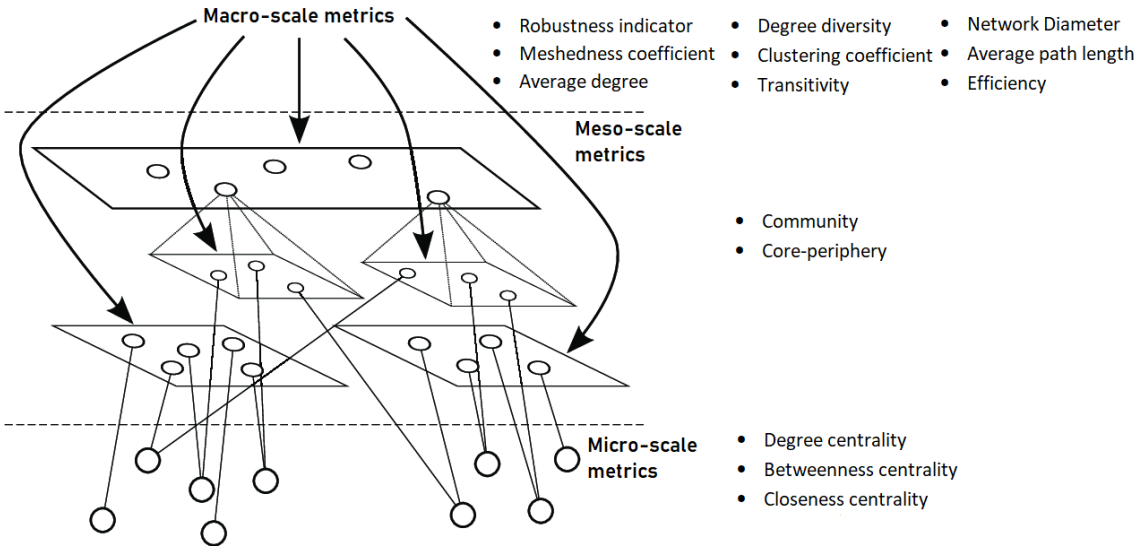


Figure 3.1: Conceptual visualisation of the micro-, meso-, and macro-scale metrics. The figure was adapted and modified from Sanderson et al. (2012).

3.1 Macro-scale Metrics

By macro-scale metrics, we are referring to metrics that describe the network as a whole. Often this is an average value of certain measures. We consider nine such global metrics. They can be further segregated into three categories based on the specific features considered by the metric.

- Macro-scale metrics based on overall network size and links. Specifically, we consider the robustness indicator and meshedness coefficient which are both computed as a function of N and L.
- Macro-scale metrics based on local connectivity. These metrics consider the average of some local features of each node. The prominent input for these metrics is based on the node degrees and/or their immediate neighbors. For this, we consider average degree, degree diversity, clustering coefficient, and transitivity.
- Macro-scale metrics based on paths. These metrics consider the paths between pairs of nodes to characterise the network. We consider network

diameter, average path length, and network efficiency for this.

In the following, we introduce all the metrics mentioned above.

3.1.1 The robustness indicator, r^T

The robustness indicator, first introduced in Derrible and Kennedy (2010), was later modified by Wang et al. (2017) and was used to measure the robustness of metro networks. In its original form, it considers the number of alternative paths in the network as a ratio of those paths over the total number of nodes in the network. It considers multi-edge scenarios where there may be multiple links between a pair of nodes. Similar to Wang et al. (2017), we also do not consider multi-edge graphs and our networks are also sufficiently sparse. Essentially, r^T increases when there are more alternative paths available to reach a destination, and decreases in larger systems that may be more difficult to maintain.

Thus, for this work, we use the following, adapted from Derrible and Kennedy (2010):

$$r^T = \frac{ln(L - N + 2)}{N}, \tag{3.1}$$

with L the number of links and N the number of nodes (intersections).

3.1.2 Meshedness coefficient, m

In the context of street networks, Cardillo et al. (2006) proposed the concept of meshedness coefficient, m, by considering the network as a planar graph since in reality, road networks are spatially embedded networks. The coefficient exploits the Euler formula, $F = L - N + 1$ as the numerator where F is the number of faces of the planar graph while the maximum number of faces when considering a maximally connected planar graph, $F_{\text{max}} = 2N - 5$ as the denominator (Wang et al. 2017):

$$m = \frac{F}{F_{\text{max}}} = \frac{L - N + 1}{2N - 5}. \tag{3.2}$$

As the name implies, meshedness coefficient is a measure that quantifies the degree of interconnectedness or 'meshedness' in the network. One property of this measure is that it ranges between 0 and 1. With maximal planar graphs where $L = 3N - 6$, we obtain $m = 1$ while $m = 0$ for tree graphs. A high meshedness coefficient indicates a high level of connectivity and redundancy in the road network (Wang et al. 2013; Lin and Ban 2013). This means that there are multiple routes or paths between different locations, which can increase the robustness and resilience of the network. Conversely, a low meshedness coefficient indicates a less connected network with fewer alternative routes (Knaap and Rey 2023).

3.1.3 Average degree, $\bar{d}$

One of the most commonly used quantities in complex network analysis is the node degree which indicates the number of immediate neighbors of a node. For instance, in the context of road networks, this has been used in various forms (Lee and Jung 2018; Zeng et al. 2020; Shang et al. 2020a) with the simplest being the average degree of all nodes in the considered network which could be obtained as follows (Diestel 2024):

$$\bar{d} = \frac{1}{N} \sum_{i=1}^{N} d_i \tag{3.3}$$

where d_i is the degree of node v_i, i.e., $d_i = \sum_j^N a_{i,j}$. A network with higher $\bar{d}$ indicates higher link density which in turn implies better network robustness.

3.1.4 Degree diversity, κ

A closely related degree-based metric of average degree is the degree diversity, the second-order average, which is given as Wang et al. (2017):

$$\kappa = \frac{\sum_{i=1}^{N} d_i^2}{\sum_{i=1}^{N} d_i}. \tag{3.4}$$

This quantity is often used in epidemic and percolation theory. For instance, in epidemic theory, it relates to the epidemic threshold below which the epidemic will

die off (Pastor-Satorras and Vespignani 2001; Boguná et al. 2003). To normalise the degree diversity within the range of $[0, 1]$, we take the inverse of the degree diversity.

3.1.5 Clustering coefficient, CC_G

First introduced in Watts and Strogatz (1998), the clustering coefficient has become a widely used metric to assess how the neighbors of a node are connected with one another. The clustering coefficient of a node represents the proportion of links that exist between its neighboring nodes out of the maximum possible number of links (Boeing 2018; Boeing 2017; Schank and Wagner 2005).

The clustering coefficient of node v_i is defined as:

$$CC_i = \frac{2E_i}{d_i(d_i - 1)}, \tag{3.5}$$

where E_i is the number of links connecting neighbors of nodes v_i and d_i is the degree of node v_i. The clustering coefficient of a node v_i characterises the connection density among the neighbors of node v_i. The maximum clustering coefficient is achieved in a complete graph where all the neighbors of a node are connected. From the above, we can then obtain the average clustering coefficient as follows (Watts and Strogatz 1998):

$$CC_G = \frac{1}{N} \sum_{i=1}^{N} CC_i. \tag{3.6}$$

For a graph with N nodes, the clustering coefficient is bounded by $0 \leq CC_G \leq 1$ where 0 is obtained in a tree and 1 is reached in a complete graph. In our context, it serves as a measure of the connectivity of a road network by assessing the degree to which the neighborhood of a specific intersection/junction is interconnected (Jiang and Claramunt 2004; Boeing 2017). The average clustering coefficient is likely to be higher in spatially embedded networks such as road networks due to the fact that nodes located in close proximity are more likely to be connected (Barthelemy 2018).

3.1.6 Transitivity, τ_G

Transitivity was introduced by Newman et al. (2002) and Karlberg (1997), as an alternative formulation of the clustering coefficient of a graph (Schank and Wagner 2005). Let G denote a simple and undirected graph. A triangle, represented by $\Delta = (\mathcal{V}_\Delta, \mathcal{E}_\Delta)$, is a complete subgraph of G consisting of exactly three nodes. The number of triangles in graph G is denoted as $\lambda(G)$, and it can be calculated by taking one-third of the sum of the triangle counts for each node, i.e., $\lambda(G) = \frac{1}{3}\sum_{i \in V} \lambda(i)$.

A triple refers to a subgraph in G with three nodes and two edges. A triple at v_i is defined as one where v_i is incident to both edges of the triple. The number of triples at v_i in terms of its degree d_i is given by:

$$T_G = \frac{d_i^2 - d_i}{2} \tag{3.7}$$

The total number of triples in the entire graph is determined by summing up the triple counts for each node, denoted as $T(G) = \sum_{i \in V} T(i)$.

The transitivity of a graph, τ_G, is defined as the ratio of three times the number of triangles in G to the number of triples in G, as proposed in Brinkmeier and Schank (2005):

$$\tau_G = \frac{3 \times \lambda(G)}{T_G} \tag{3.8}$$

Since each triangle contains exactly three triples, it follows that $3\lambda(G) \leq T(G)$. Consequently, the transitivity $\tau(G)$ is always a rational number ranging from 0 to 1. Transitivity is closely related to the clustering coefficient and sometimes used as an alternative where both provide insights into the local connectivity structure of a network (Schank and Wagner 2005; Chehreghani and Chehreghani 2014).

3.1.7 Network diameter, D

Distance between nodes in a network is often important in network analysis. This is also true for road networks. The network diameter is one of the most basic quantities describing the network in terms of node distance. It is the maximum node eccentricity that corresponds to the longest shortest path amongst all pairs of nodes in the network of interest (Wasserman and Faust 1994).

$$D = \max_{i,j}(H_{ij}) \tag{3.9}$$

where H_{ij} denotes the shortest path distance (in hopcount) between node v_i and v_j. When G is not connected, we find the diameter of each connected component and take the maximum value.

3.1.8 Average path length, H_G

While network diameter only uses one particular path as an indicator of the graph property, a more informative measure considering all shortest paths is the average path length, H_G (Newman 2012). Essentially, it is the mean of shortest path lengths between all node pairs (Newman et al. 2002):

$$H_G = \frac{2}{N(N-1)} \sum_{i=1}^{N} \sum_{j=1}^{N} H_{i,j} \tag{3.10}$$

where N is the number of nodes in the network.

3.1.9 Efficiency, eff_G

The measure of network efficiency provides an indication of how easily traffic spreads through an arbitrary network. It is given by the average of reciprocals of all the shortest paths between node pairs and can be given as follows (Latora and Marchiori 2002):

$$eff_G = \frac{2}{N(N-1)} \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{1}{H_{i,j}}. \tag{3.11}$$

In this measure, the network is considered to be more efficient when nodes can reach each other in fewer hops (i.e., shorter paths).

3.2 Meso-scale Metrics

Mesoscopic structures relating to subgraph patterns in complex networks have recently gained much attention. In this work, we focus on two main concepts: community and core-periphery structures.

3.2.1 Community

A population often consists of different communities where a subset of nodes have more connections among them than to others. A string of research has attempted to capture such phenomena. Given the social contact network, the goal is to find community structures in the graph. In this work, we exploit the concept of *modularity*, Q, one of the most used functions in determining communities in a network (Newman 2004), which is measured by using the following:

$$Q = \frac{1}{2L} \sum_{i,j} (a_{ij} - \frac{d_i d_j}{2L}) \mu(c_i, c_j), \tag{3.12}$$

where c_i is the community of i, c_j that of j, the sum goes over all i and j pairs of nodes, and $\mu(c_i, c_j)$ is 1 if $c_i = c_j$ and 0 otherwise. Various modularity algorithms have been proposed in the literature, and in our work, we use the Clauset-Newman-Moore algorithm (Clauset et al. 2004)[1].

3.2.2 Core-periphery

Another meso-scale structure that is often found in real networks, such as social, neural, and transportation networks, is a core-periphery structure where nodes in the network can be categorised as core which is densely inter-connected between them, and periphery nodes which are adjacent to core nodes but not to each

[1] Modularity maximisation is an NP-complete problem (Brandes et al. 2006). It is beyond the scope of this research to investigate the goodness of the various available algorithms.

other (Borgatti and Everett 2000). In this work, we have employed the transport-based core-periphery detection algorithm, referred to as path-core *PC* algorithm, proposed in Cucuringu et al. (2016) which is suitably based on paths rather than immediate neighbours[2].

According to the *PC* algorithm, core-periphery structure identification bears a resemblance to betweenness centrality (defined later in Section 3.3.2) in network analysis (Cucuringu et al. 2016).

The Path-Core metric of node v_i, $PC(v_i)$, is defined as:

$$PC(v_i) = \sum_{\substack{(j,k)\in E \\ (j,k)\neq v_i}} \frac{\sigma_{jk}(v_i)|_{G\setminus(j,k)}}{\sigma_{jk}|_{G\setminus(j,k)}} \tag{3.13}$$

where, $\sigma_{jk}|G \setminus (j,k)$ counts the number of shortest paths between node j and k in the network G after removing the edge (j,k) itself. Additionally, $\sigma_{jk}(v_i)|G \setminus (j,k)$ counts the subset of these paths that pass through node v_i.

3.3 Micro-scale Metrics

For micro-scale metrics, we are looking into nodal-level measures of nodes in a network. Specifically, we exploit widely used centrality measures to rank nodes or links based on their relative importance and examine their distributions to derive insights into the characteristics of given networks. Various works, such as Crucitti et al. (2006), Lämmer et al. (2006), and Barthélemy (2011) have already utilised centrality measures in the context of road networks, and in this work, we choose the three most commonly used centrality measures as follows:

3.3.1 Degree centrality

The simplest yet most widely used centrality measure is the degree centrality. The degree centrality of a node is the number of direct neighbors the node has

[2]We have also considered different algorithms (e.g., Borgatti's algorithm (Borgatti and Everett 2000)) but most proved to be computationally infeasible for the large datasets we consider here.

(see Section 3.1.3). Networks with different degree distributions can exhibit vastly different properties. For instance, Erdös-Rényi random graph model (Erdos and Renyi 1959) has a binomial degree distribution (or a Poisson distribution in the limit of large N) while scale-free networks (Barabasi 2013) has power-law degree distribution and both models has different properties (Erdos and Renyi 1959).

Degree centrality of a node v_i, $c_D(v_i)$, is defined as follows:

$$c_D(v_i) = \frac{d_i}{N-1} \tag{3.14}$$

3.3.2 Betweenness centrality

The betweenness centrality (Wasserman and Faust 1994) measures the number of times a node lies on the shortest paths between all pairs of nodes in a network. It appears as a natural representation of *load* in terms of traffic in road networks. With the assumption that road users normally use the shortest route to get to their destinations, then betweenness centrality provides a good indicator to measure the burden of nodes in the transport process. Even in instances where movement within the network is random, locations with high betweenness centrality experience higher volumes of traffic or congestion, as indicated by Crucitti et al. (2006) and Barthélemy (2011). Hence, nodes characterised by high betweenness centrality are more susceptible to disruptions and congestion.

Betweenness centrality of a node v_i, $c_B(v_i)$, is defined as follows:

$$c_B(v_i) = \sum_{v_j, v_k \in \mathcal{V}} \frac{\sigma(v_j, v_k|v_i)}{\sigma(v_j, v_k)}, \tag{3.15}$$

where $\sigma(v_j, v_k)$ is the number of shortest paths between v_j and v_k and $\sigma(v_j, v_k|v_i)$ is the number of those paths passing through v_i.

Betweenness distribution has been found to follow power-law in scale-free graphs which are used to model various types of real-world networks such as the Internet and road networks (Wang et al. 2008). Such betweenness distribution is exploited in path-based problems (e.g., in content caching problems in computer networks

(Chai et al. 2013)).

3.3.3 Closeness centrality

Closeness centrality (Wasserman and Faust 1994) measures how close a node, in terms of shortest path lengths, to all other nodes in the network. It is defined as the reciprocal of the average of those $N-1$ shortest paths and can be computed as follows:

$$c_c(v_i) = \frac{N}{\sum_{j=1;j\neq i}^{N-1} H_{i,j}}. \tag{3.16}$$

While betweenness centrality focuses on the role of a node as a connector or intermediary between other nodes, closeness centrality focuses on the distance of a node to other nodes in the network. A node that is closer to all others is deemed to be of higher relative importance (e.g., a higher level of accessibility with others).

The macroscopic, mesoscopic, and microscopic measures discussed earlier will be elaborated upon in the next sub-section, detailing the traffic data employed in this study.

3.4 Road Network Dataset

We consider 29 road networks from China and India, the top two most populous countries in the world. Specifically, 20 Chinese and 9 Indian cities are selected for this work. Their locations are shown in Figure 3.2, the higher the population of a city, the deeper the colour of its corresponding node.

We use the dataset provided in Karduni et al. (2016), who have developed a tool called GIS Features 2 Edgelist (GISF2E). This tool allows for the conversion of shapefiles into networks and was applied by Karduni et al. (2016) to the road networks of 80 highly populated cities worldwide, using data from OpenStreetMap. The extracted data includes the geospatial coordinates of nodes, the length of each road segment, and information on which nodes are connected by each link

within the network. The geographic coordinates are represented in Universal Transverse Mercator format, while the link lengths are provided in meters. With this data, a network can be constructed, as each observation in the dataset represents an edge, indicating the two connected nodes, their coordinates, and the link length in meters. Map of the studied 29 cities are visualised in Figure 3.4 and Figure 3.5.

Table 3.1 lists the basic statistics of the dataset for the cities considered in this work. The size of the studied road networks varies, ranging from $2,593$ nodes to $151,947$ nodes. A preliminary examination of the table reveals a strong linear correlation between the number of nodes and links for the chosen road networks as shown in Figure 3.3 with $R^2 = 0.99$ for the linear fit where we obtain $L = 2.62N + 1797.23$.

Figure 3.2: Studied cities: from less (light red) to more (black) populated (personal collection 2024).

Table 3.1: Size of studied road networks

Road networks	Country	Number of Nodes, N	Number of Edges, L	Population (million)
Surat	India	2593	7340	7.86
Quanzhou	China	5672	15234	1.83
Dongguan	China	8315	22256	7.52
Zhengzhou	China	9162	25730	5.74
Harbin	China	10727	29422	6.7
Fuzhou	China	12333	32338	3.86
Ahmedabad	India	12859	36406	8.5
Shenyang	China	13000	38052	7.57
Dalian	China	13605	35794	5.92
Qingdao	China	13894	38036	5.89
Chengdu	China	18300	51860	9.52
Hangzhou	China	20745	56682	8.1
Nanjing	China	21248	59070	9.5
Wuhan	China	21560	60016	8.62
Chongqing	China	21779	54198	17.01
Xianyang	China	26587	73114	8.61
Pune	India	28649	73368	7.03
Tianjin	China	31696	86186	14.07
Shenzhen	China	37004	99908	12.89
Suzhou	China	46094	124380	1.28
Mumbai	India	46660	121960	21.04
Guangzhou	China	52043	138760	14.05
Delhi	India	53689	151970	32.31
Chennai	India	54440	146344	11.57
Calcutta	India	71554	179424	15.17
Shanghai	China	78560	213456	28.71
Beijing	China	83884	220256	21.45
Hyderabad	India	130414	343954	10.6
Bengaluru	India	151947	398956	13.31

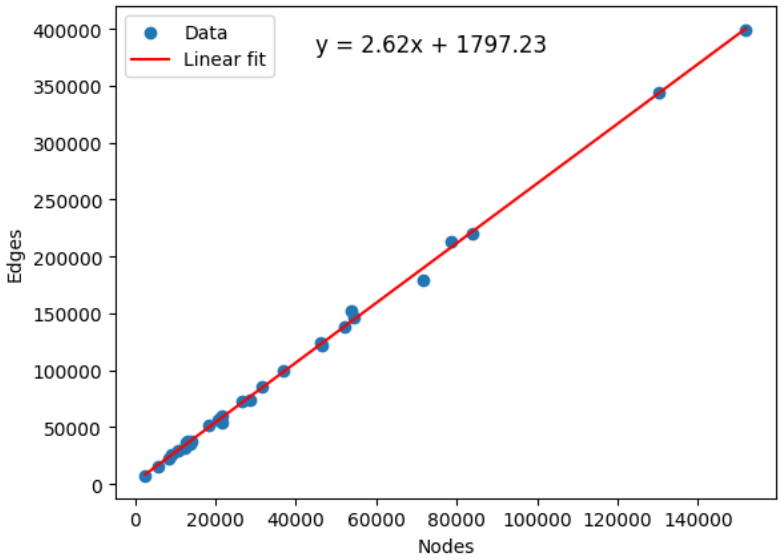


Figure 3.3: Linear relationship graph between the number of nodes and edges in studied 29 road networks.

3.5 Results and Analysis

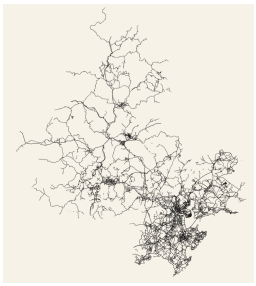
In this section, we will share the outcomes of evaluating metrics at different levels of scale: macro, meso, and micro.

3.5.1 Assessment at macro-scale

We first present the results of the macro-scale metrics in Table 3.2 and Table 3.3. The robustness indicator, r^T, suggests that Surat, Dongguan, and Zhengzhou have the most robust road networks among the 29 considered cities. From our results, it appears r^T favors smaller networks where the robustness has an inverse relationship with the size of the network. We do not observe such a relationship with the meshedness coefficient, m, even though similar to r^T, both metrics are a function of L and N. A higher m is generally considered desirable for road networks as it implies better accessibility, flexibility, and resilience in terms of transportation options (Wang et al. 2013; Lin and Ban 2013). In this case, Shenyang has the highest meshedness which also has the highest average degree, $\bar{d}$.

However, we see a small deviation of $\bar{d}$ across the 29 cities (all having $\bar{d}$ between

(a) Surat

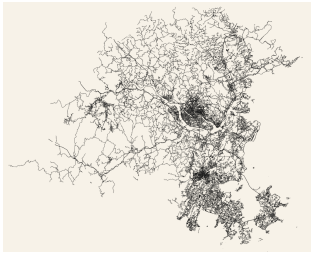


(b) Quanzhou

(c) Dongguan

(d) Zhengzhou

(e) Harbin

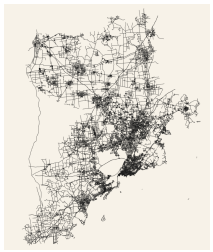


(f) Fuzhou

(g) Ahmedabad

(h) Shenyang

(i) Dalian



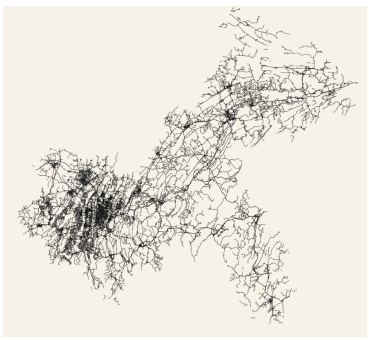
(j) Qingdao

(k) Chengdu

(l) Hangzhou

(m) Nanjing

(n) Wuhan



(o) Chongqing

Figure 3.4: Maps of chosen cities for this study and obtained from an open-source visualisation web (Anvaka GitHub 2024).

(a) Xianyang

(b) Pune

(c) Tianjin

(d) Shenzhen

(e) Suzhou

(f) Mumbai

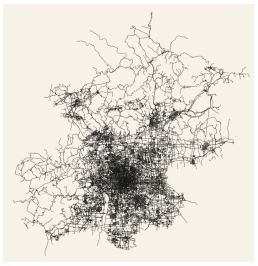
(g) Guangzhou

(h) Delhi

(i) Chennai

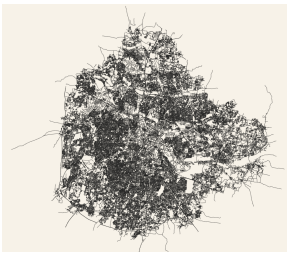
(j) Calcutta

(k) Shanghai



(l) Beijing

(m) Hyderabad



(n) Bengaluru

Figure 3.5: Maps of chosen cities for this study (continued) and obtained from an open-source visualisation web (Anvaka GitHub 2024).

2.5 to 2.9). This corroborates with previous findings (e.g., Reza et al. 2022; Lee and Jung 2018; Akbarzadeh et al. 2018) that urban road networks do not exhibit large variation in degrees due to planar constraints (Lämmer et al. 2006; Viana et al. 2013). In our case, we see that the densely populated cities have road networks having most junctions connecting two or three road segments. As an artifact of this, we also note that the degree diversity, κ, also exhibits small deviations and positively correlates with $\bar{d}$.

While degree-based metrics such as $\bar{d}$ and κ focus on the number of immediate neighbors, clustering coefficient, CC_G, and transitivity, τ_G, measure how well these immediate neighbors are connected to each other (i.e., the triangles in a network). They differ in their method of sampling (see Rohe 2013). In our context, they assess how well the neighboring nodes of a road junction or road intersection are interconnected between them. For instance, a high clustering coefficient indicates a network that is highly clustered, while a low clustering coefficient indicates a sparsely connected network. From our results, the two metrics provided different but still correlated rankings (e.g., Wuhan and Quanzhou are highest in CC_G but Beijing and Fuzhou are highest in τ_G).

The three remaining macro-scale metrics (i.e., D, H_G, and eff_G) are all based on shortest paths. We note that both the diameter and average shortest paths of all the chosen networks are relatively high (at an average of 198 hops and 68 hops respectively from the 29 cities). This is due to the level of abstraction in the dataset where every junction and intersection is extracted as a node and particularly in big cities considered here, the number of nodes becomes large even across small spatial areas. An inter-city road network (where cities are nodes) will record much smaller D and H_G due to the different methods of abstraction (in this case, cities are nodes). This then points to the observation that these urban city road networks do not exhibit small-world property (Watts and Strogatz 1998). Also resulting from this, the network efficiency for the road networks is low (since by definition, efficiency is computed as a function of the inverse of shortest path lengths).

From the above discussion, we can already infer some relationships between the

Table 3.2: Macro-scale metrics results.

Road networks	r^T	m	$\bar{d}$	κ	CC_G	τ_G	D	H_G	eff_G
Surat	0.00326	0.916	2.8	3.10	0.048	0.057	73	29	0.044
Quanzhou	0.00162	0.843	2.7	3.09	0.053	0.079	125	43	0.028
Dongguan	0.00115	0.839	2.7	3.05	0.036	0.049	135	46	0.026
Zhengzhou	0.00106	0.904	2.8	3.13	0.03	0.04	114	42	0.030
Harbin	0.00092	0.872	2.7	3.09	0.039	0.05	162	53	0.024
Fuzhou	0.00080	0.811	2.6	3.01	0.050	0.08	128	46	0.021
Ahmedabad	0.00078	0.916	2.8	3.11	0.039	0.047	129	49	0.026
Shenyang	0.00078	0.964	2.9	3.23	0.050	0.07	117	41	0.029
Dalian	0.00074	0.816	2.6	3.01	0.046	0.06	186	70	0.018
Qingdao	0.00073	0.869	2.7	3.15	0.030	0.04	183	63	0.021
Chengdu	0.00057	0.917	2.8	3.19	0.040	0.06	137	51	0.024
Hangzhou	0.00051	0.866	2.7	3.11	0.040	0.059	201	62	0.02
Nanjing	0.00050	0.890	2.8	3.12	0.047	0.06	162	56	0.021
Wuhan	0.00049	0.892	2.8	3.11	0.056	0.073	143	52	0.023
Chongqing	0.00048	0.744	2.5	2.81	0.034	0.05	439	105	0.014

Table 3.3: Macro-scale metrics results (continued).

Road networks	r^T	m	$\bar{d}$	κ	CC_G	τ_G	D	H_G	eff_G
Xianyang	0.00040	0.875	2.8	3.08	0.036	0.05	223	63	0.019
Pune	0.00037	0.781	2.6	2.91	0.029	0.038	170	68	0.018
Tianjin	0.00034	0.860	2.7	3.09	0.032	0.04	186	66	0.018
Shenzhen	0.00030	0.850	2.7	3.08	0.043	0.06	190	65	0.017
Suzhou	0.00024	0.849	2.7	3.06	0.037	0.06	257	91	0.013
Mumbai	0.00024	0.807	2.6	2.98	0.04	0.058	325	106	0.012
Guangzhou	0.00022	0.833	2.7	3.02	0.040	0.055	226	78	0.015
Delhi	0.00021	0.915	2.8	3.09	0.029	0.030	190	76	0.016
Chennai	0.00021	0.844	2.7	2.99	0.023	0.030	227	79	0.016
Calcutta	0.00016	0.754	2.5	2.87	0.021	0.029	303	101	0.013
Shanghai	0.00015	0.859	2.7	3.07	0.028	0.04	263	86	0.014
Beijing	0.00014	0.813	2.6	3.11	0.040	0.1	244	90	0.011
Hyderabad	0.00009	0.819	2.6	3.02	0.024	0.030	267	108	0.011
Bengaluru	0.00009	0.813	2.6	2.97	0.029	0.037	254	102	0.011

macro-scale metrics. We present the Pearson correlations of the nine metrics in Figure 3.6.

From the figure, we note the strong positive correlations among m, $\bar{d}$ and κ where we obtain $\rho(m, \bar{d}) = 0.97$, $\rho(m, \kappa) = 0.89$ and $\rho(\bar{d}, \kappa) = 0.84$. They are consistent (e.g., Shenyang has the highest value for all three metrics: m=0.964, $\bar{d}$=2.9, and κ=3.23). Clustering coefficient and transitivity appear to be another related group of metrics, unsurprisingly due to them measuring similar features of the network ($\rho(CC_G, \tau_G) = 0.82$). With both taking the shortest path as the input, diameter and average shortest path lengths also exhibit high correlation ($\rho(D, H_G) = 0.91$).

Interestingly, we computed $\rho(r^T, eff_G) = 0.90$, implying those networks with high robustness indicators are also the most efficient ones. On the other hand, H_G and eff_G show significant negative relation ($\rho(H_G, eff_G) = -0.89$). The same tendency was found also in Shang et al. (2020a) on nine smaller road networks. In general, we found that both D and H_G have a negative correlation with other metrics.

Finally, we summarise our results in radar diagrams in Figure 3.7 and Figure 3.8, providing a visual illustration of the various considered metrics. For this purpose, each value of x_i in a set of macro-scale metrics from Table 3.2 and Table 3.3 is re-scaled to a $x_{i_{new}}$ value within the range [0,1], by normalisation formula: $x_{i_{new}} = \frac{x_i - x_{min}}{x_{max} - x_{min}}$. From the diagrams, we note that smaller road networks achieve higher values in macro-metrics at the upper half of the radars (i.e., Surat, Quanzhou, Zhengzhou, Ahmedabad, Shenyang, Qingdao, Chengdu, Hangzhou, Nanjing and Wuhan).

On the other hand, bigger road networks tend to achieve higher values in the lower half of the radar diagrams (i.e., Chongqing, Suzhou, Mumbai, Guangzhou, Calcutta, Hyderabad, and Bengaluru). We show the area of the shaded polygon in the radar diagrams. Shenyang and Surat obtain higher values in more metrics while at the other end of the spectrum, Calcutta and Pune achieve the lowest overall area.

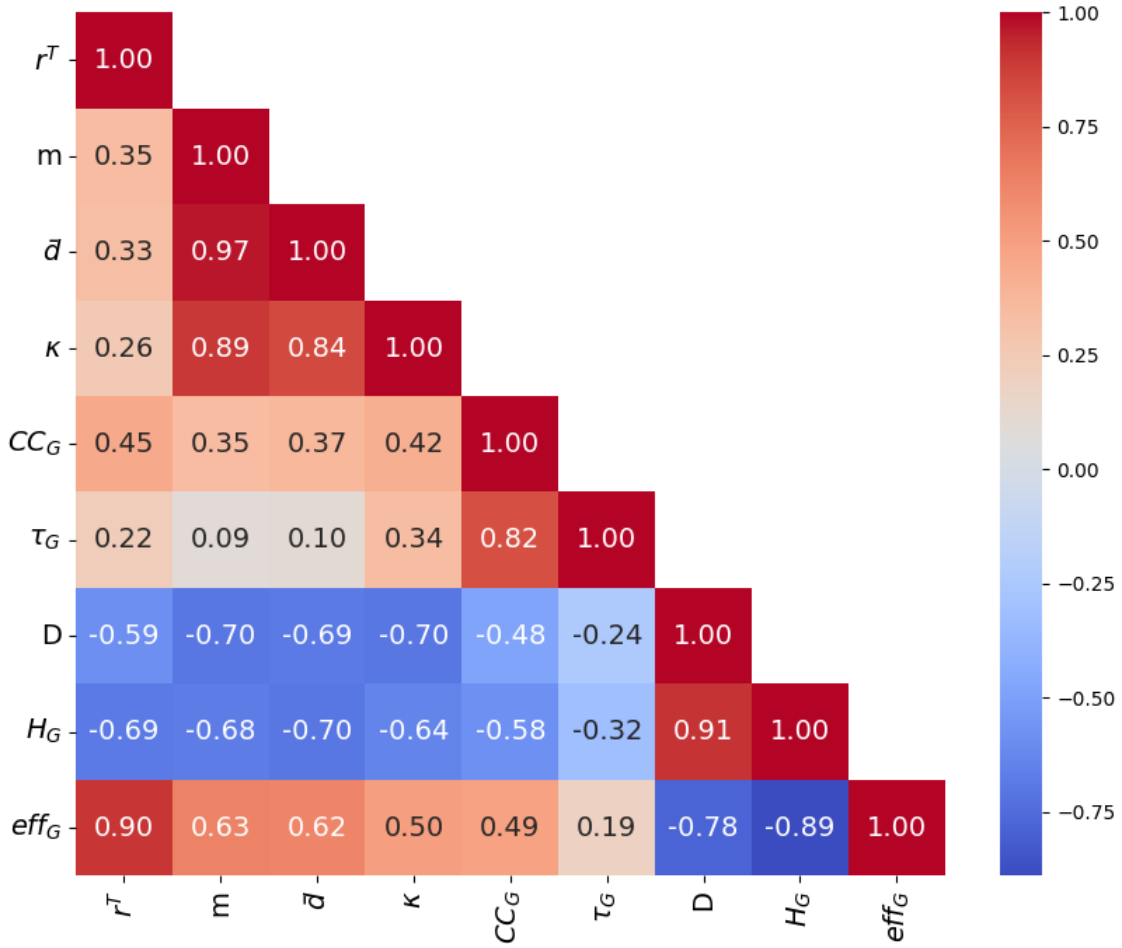


Figure 3.6: Pearson correlation heatmap of macro-scale metrics

3.5.2 Assessment at meso-scale

To detect meso-scale structures, we compute the modularity, Q, number of communities, and average community size for community structure while for core-periphery structure, we compute the Path-Core (PC)%. These are presented in Table 3.4. The first observation must be the unusually high modularity we obtain for all cities (Q close to 1.0). As indicated in Newman (2004), the usual range of modularity values typically falls between 0.3 and 0.7 and rarely achieves higher values. It appears that such densely populated urban road networks under our study represent a class of networks that exhibit consistently high modularity. This is important when modeling such real-world networks. The relatively high number of communities also suggests that these cities have multiple regions of high concentration (e.g., areas with high population density or regions with high ag-

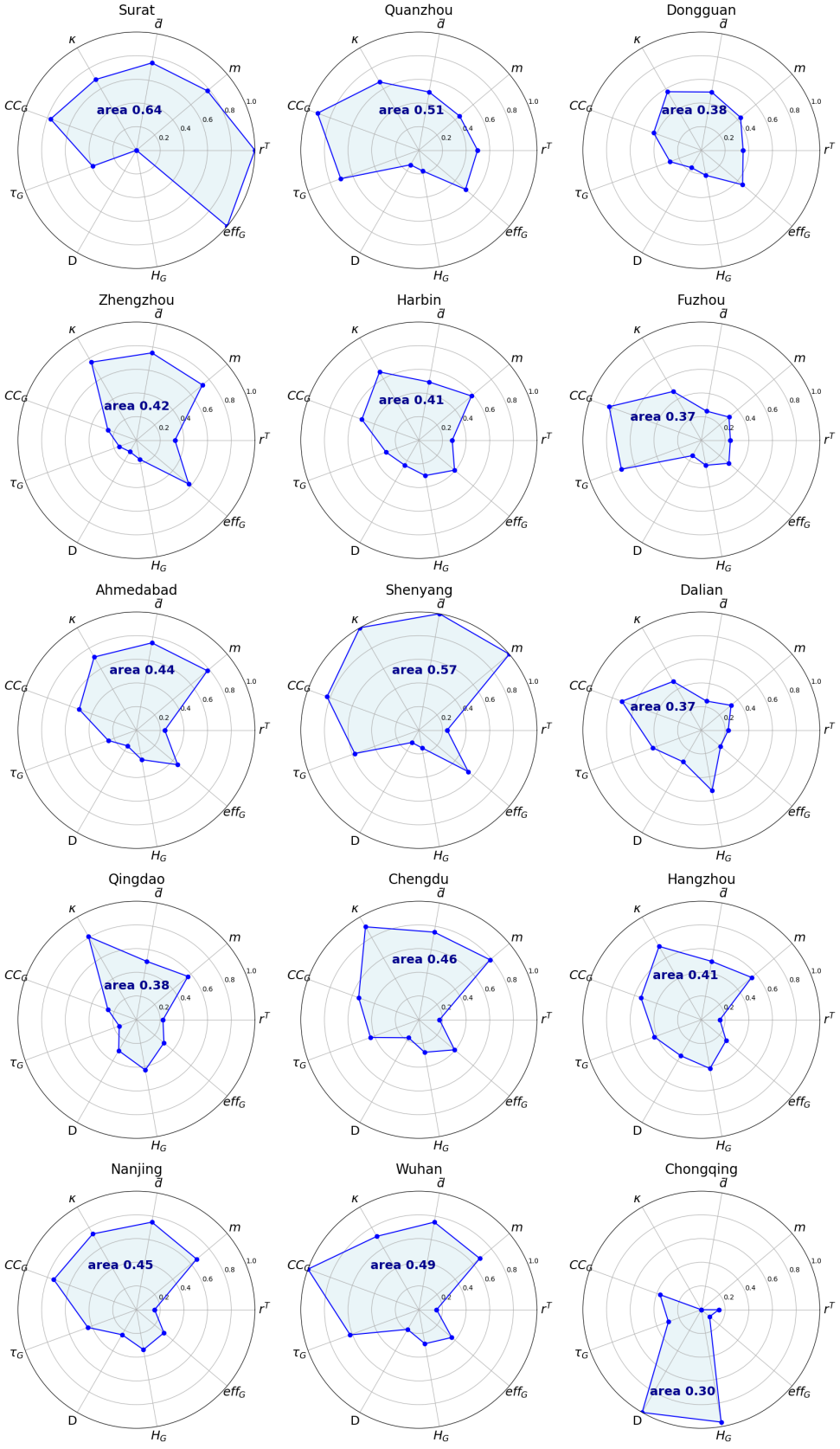


Figure 3.7: Radar diagrams with macro-scale metrics for 29 chosen cities.

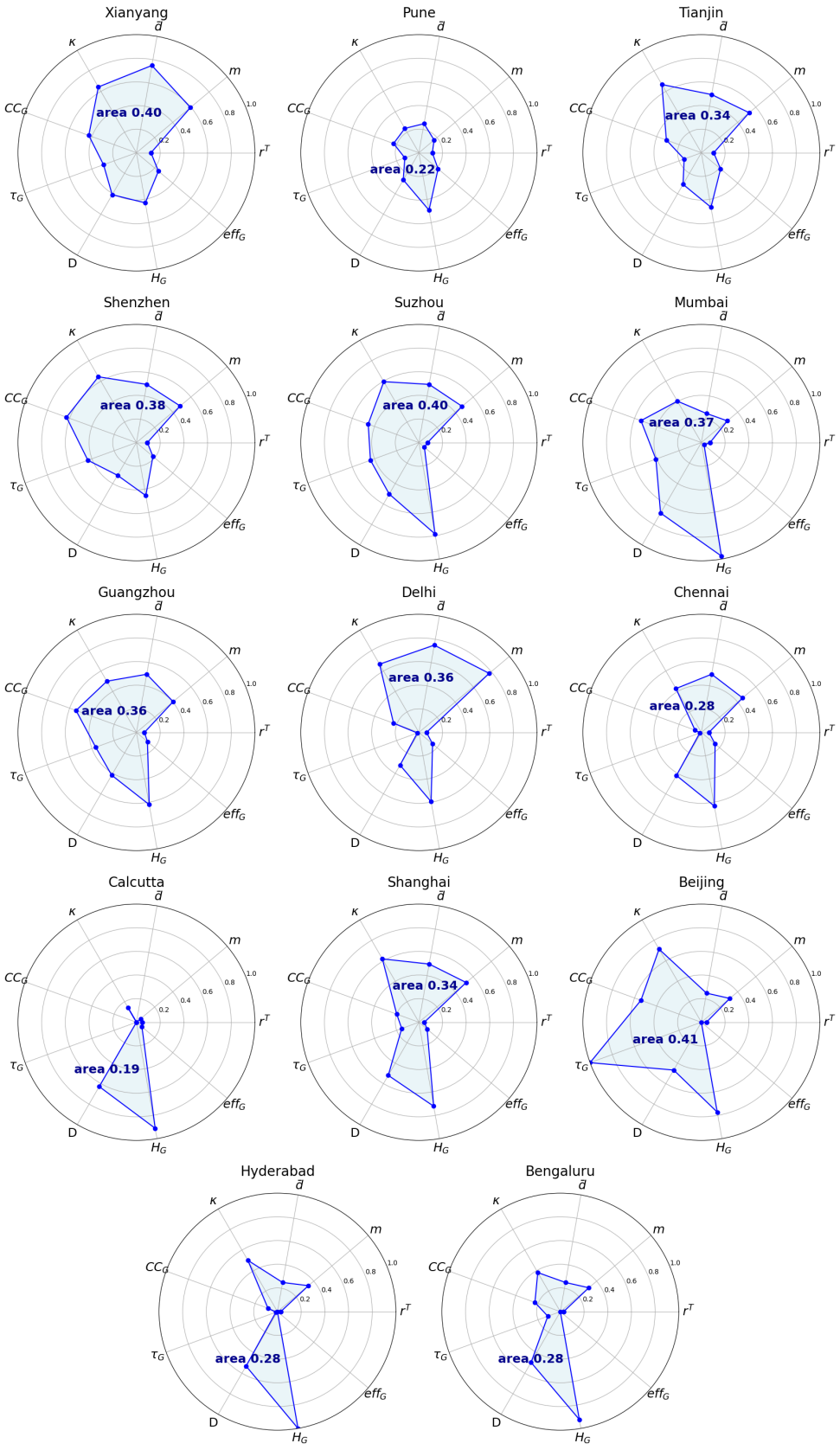
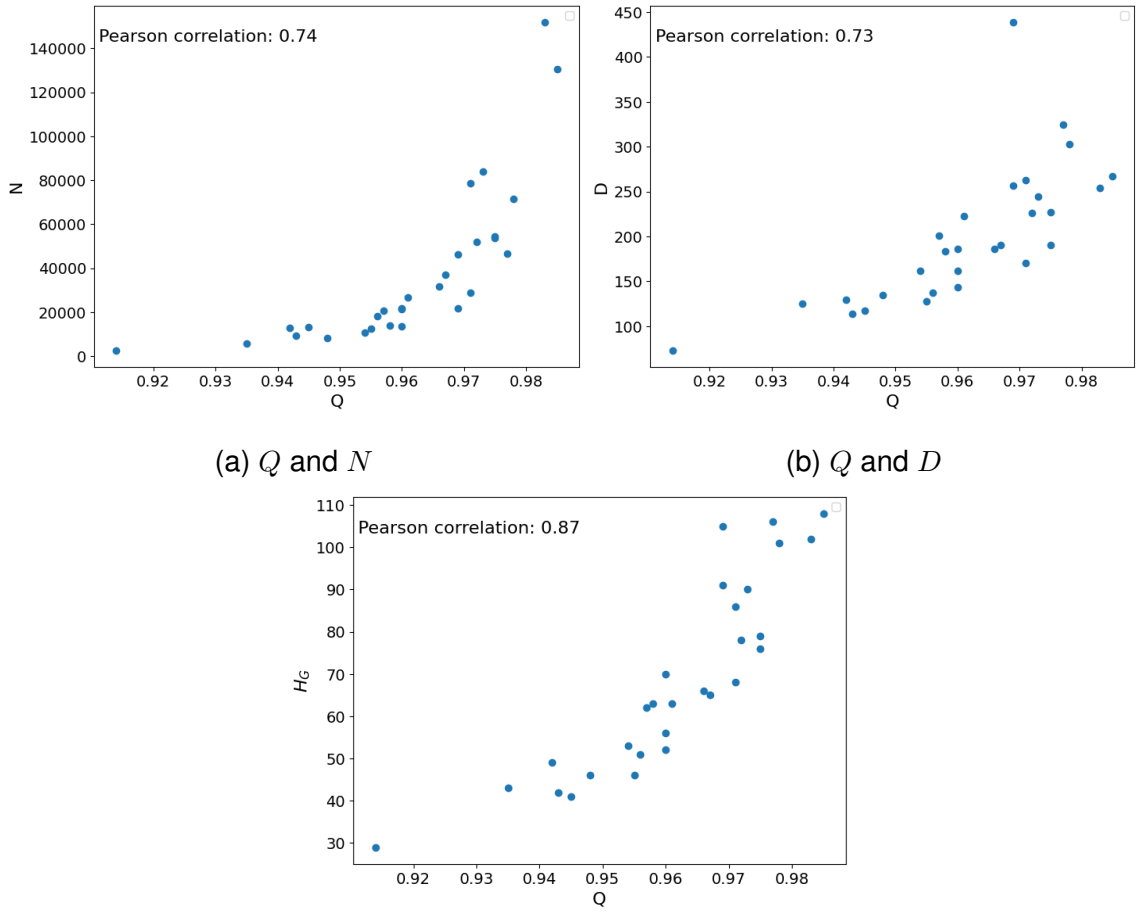


Figure 3.8: Radar diagrams with macro-scale metrics for 29 chosen cities (continued).

gregation of activities such as business districts or entertainment centres). While smaller towns may largely focus on the city centre, our results suggest that these populous cities appear to be polycentric (Kloosterman and Musterd 2001).

In our analysis, we also found correlations between macro- and meso-scale metrics. Specifically, we find that there modularity, Q, is positively correlated with N, H_G and D as shown in Figure 3.9 (i.e. $\rho(Q, N) = 0.74$, $\rho(Q, D) = 0.73$, and $\rho(Q, H_G) = 0.87$).



(a) Q and N

(b) Q and D

(c) Q and H_G

Figure 3.9: Pearson correlation of modularity Q with D, H_G, and N.

Results of the core-periphery identification using *PC* algorithm Cucuringu et al. (2016) are presented in the last column of the Table 3.4 as the percentage of *core* nodes in each road network. Chongqing is revealed as the city with the highest percentage of core nodes at 18.7% while the lowest is Ahmedabad with only 0.45% of nodes appearing as core. Nevertheless, across all the cities, the

small percentage of nodes considered as core further reinforces the previous observation that these cities do not possess one single focal point but rather possess many areas of concentration, which may be due to multiple waves of development over time as the cities grow.

3.5.3 Assessment at micro-scale

The importance of different nodes in a network could be ranked using different centrality measures and their distributions have been widely used to describe the properties of the networks under study. As prior mentioned, we have chosen the three most used centrality measures in the literature for our study here. We begin with the degree distribution, presented in Figure 3.10 and Figure 3.11 for all the cities.

The plots reveal a lack of diverse degree values with the most common degree across all road networks being three, which corresponds to T-shaped intersections. This is in line with previous findings such as in Lee and Jung (2018) where the authors similarly highlighted that junctions with three connections are more frequently found than others and in Badhrudeen et al. (2022) which also corroborated this when studying 22 Korean urban road networks.

Furthermore, we observe that crossroad junctions (two roads intersect resulting in a degree of four) are the second highest occurrences for many of the networks, especially for smaller ones, i.e., Quanzhou, Zhengzhou, Ahmedabad, and Shenyang. On the other hand, larger city road networks (e.g., Bengaluru, Hyderabad, Beijing, Calcutta, Chennai, Delhi, and Pune) have many leaf nodes (degree = 1), likely because they are richly connected to other cities. It is also obvious that these road networks in densely populated cities do not exhibit scale-free properties as their degree distributions do not follow the power law; we obtain R^2 values ranging from 0.02 to 0.18. Similar power-law fitting experiments were conducted by Reza et al. (2022) and Akbarzadeh et al. (2018) on the road networks of Ingolstadt in Germany, Porto in Portugal, and other eight urban road networks across the world, and the outcomes aligned with those observed in the road networks of these highly populated cities.

Table 3.4: Meso-scale metrics results

Road networks	Modularity, Q	Number of communities	Mean community size	*Path-Core*%
Surat	0.914	44	59	1.35
Quanzhou	0.935	206	28	9.78
Dongguan	0.948	183	45	7.82
Zhengzhou	0.943	81	113	8.30
Harbin	0.954	190	56	5.85
Fuzhou	0.955	289	43	8.12
Ahmedabad	0.942	92	140	0.45
Shenyang	0.945	116	112	5.56
Dalian	0.960	459	30	7.63
Qingdao	0.958	274	51	6.46
Chengdu	0.956	230	80	6.69
Hangzhou	0.957	282	74	8.78
Nanjing	0.960	332	64	7.74
Wuhan	0.960	182	118	6.14
Chongqing	0.969	462	47	18.7
Xianyang	0.961	261	102	7.06
Pune	0.971	268	107	6.31
Tianjin	0.966	273	116	10.4
Shenzhen	0.967	770	48	4.63
Suzhou	0.969	364	127	12.9
Mumbai	0.977	625	75	5.19
Guangzhou	0.972	526	99	8.48
Delhi	0.975	304	177	2.58
Chennai	0.975	326	167	3.78
Calcutta	0.978	547	131	7.46
Shanghai	0.971	691	114	11.2
Beijing	0.973	3214	26	7.98
Hyderabad	0.985	910	143	2.85
Bengaluru	0.983	900	169	5.84

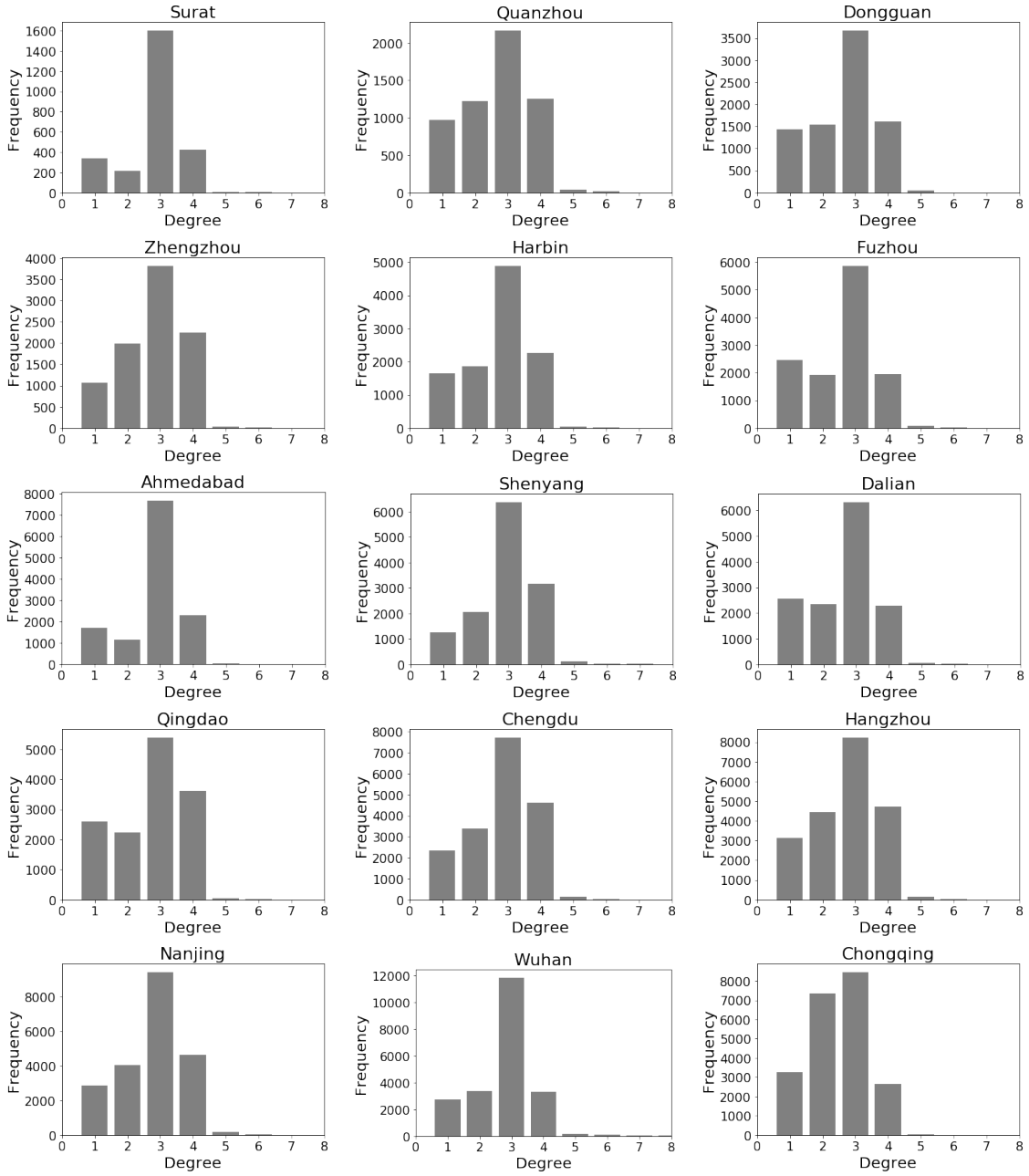


Figure 3.10: Degree distribution (personal collection 2024).

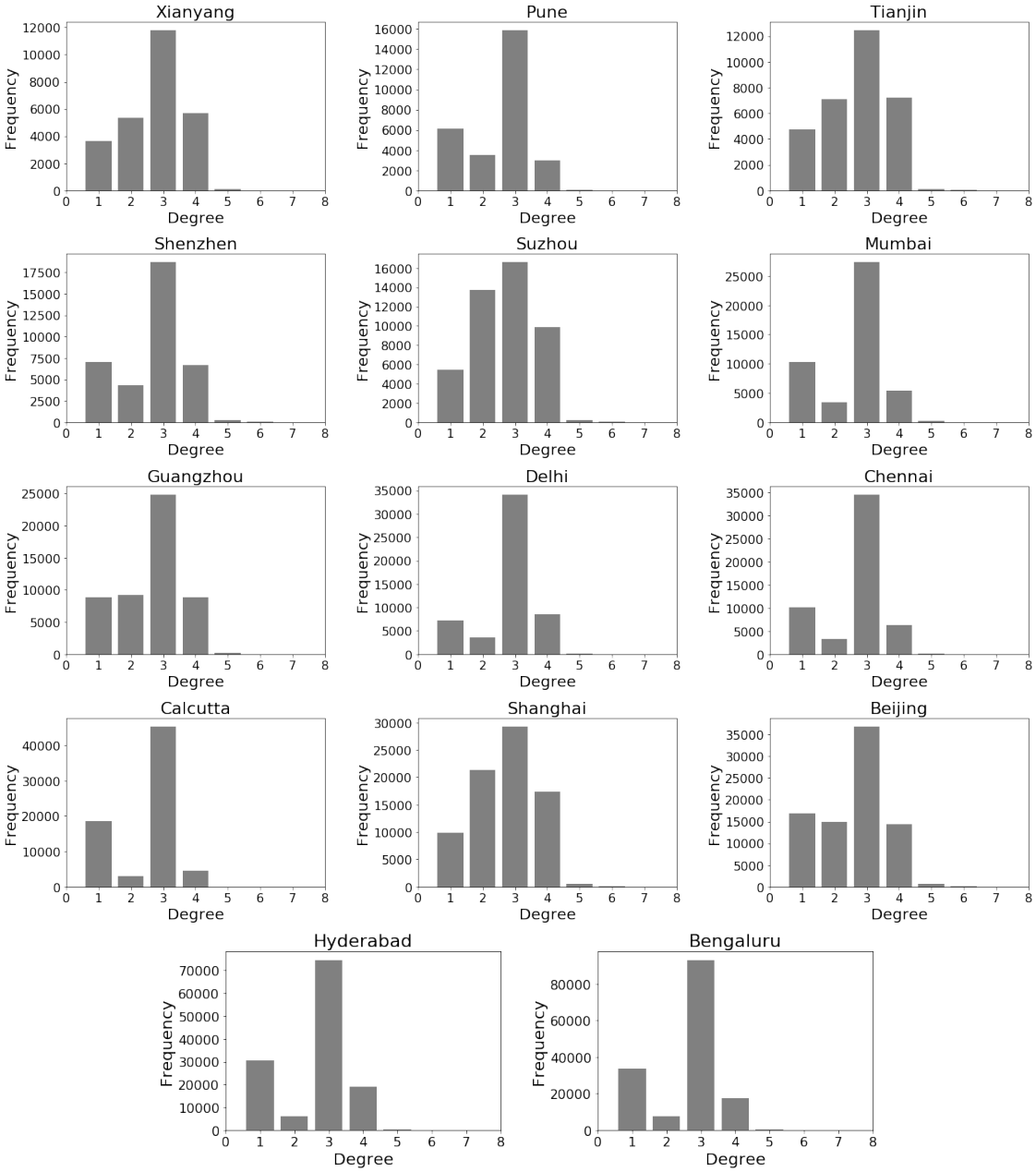


Figure 3.11: Degree distribution (personal collection 2024) (continued).

Contrary to degree distributions, we find that the betweenness distributions for the considered road networks strongly follow a power-law distribution, $P(c_B) \sim c_B^{-\beta}$ where β is the power-law exponent. Goh et al. (2001) demonstrated that networks with a power law degree distribution also exhibit a power law distribution for betweenness centrality. In the context of urban road networks, we illustrate that irrespective of the degree distribution, betweenness centrality displays a power law distribution. We show the betweenness distributions in Figure 3.12 and Figure 3.13. From our fitting, we find that the power-law exponent, β, ranges between 0.98 (Surat) and 1.39 (Calcutta) (see Figure 3.12 and Figure 3.13 for the exponent for an individual city). We also find high goodness-of-fit with all R-squared (R^2) close to 1.0.

From the perspective of road network structure, a higher value of β was interpreted as having a high concentration of traffic on the most important intersections (Lämmer et al. 2006). Investigations by Lämmer et al. (2006) and Kirkley et al. (2018) identified scale-free and truncated power-law betweenness distributions in German cities and other 97 cities worldwide. However, Crucitti et al. (2006) observes an exponential in Venice or Gaussian distribution in Richmond and San Francisco. They argue a uniform scale distribution in self-organised and planned cities, with the former displaying an exponential pattern and the latter showing a Gaussian pattern. Another study of betweenness distribution in Zurich by Casali and Heinimann (2019) exhibits a closer similarity to a log-normal distribution. Our findings align with Lämmer et al. (2006) and underline that betweenness distribution for cities in highly populated countries follows a power law.

Edge betweenness centrality c_{eB} distribution reveals a similar power law fit. This is proved by R^2 value ranging from 0.98 to 1. When the quality of fit is R^2=1, it indicates exact power-law dependence. Lämmer et al. (2006) made the same observation and concluded that there is a significant concentration of traffic volume on a small number of roads.

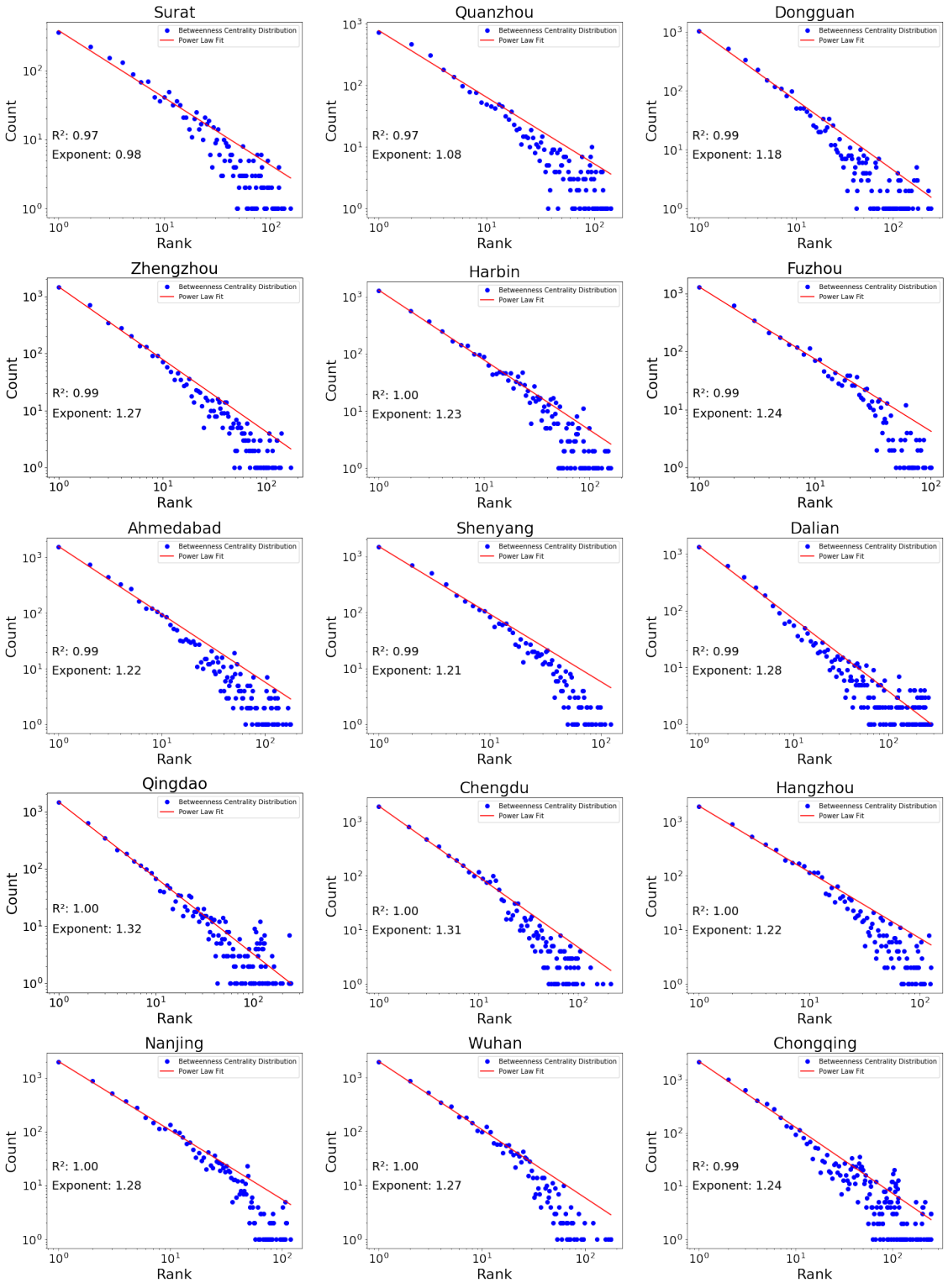


Figure 3.12: Betweenness centrality distribution (personal collection 2024).

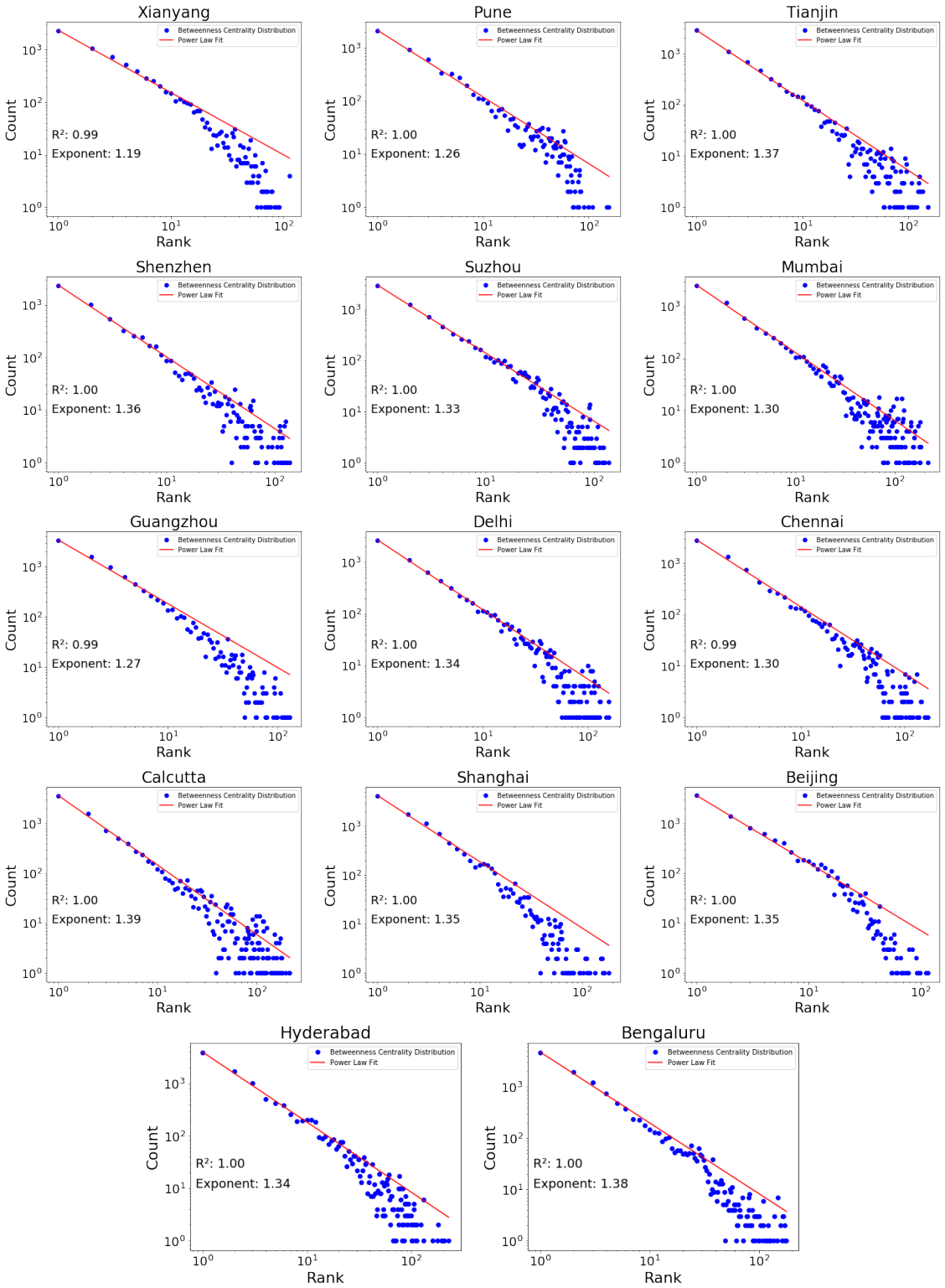


Figure 3.13: Betweenness centrality distribution (personal collection 2024) (continued).

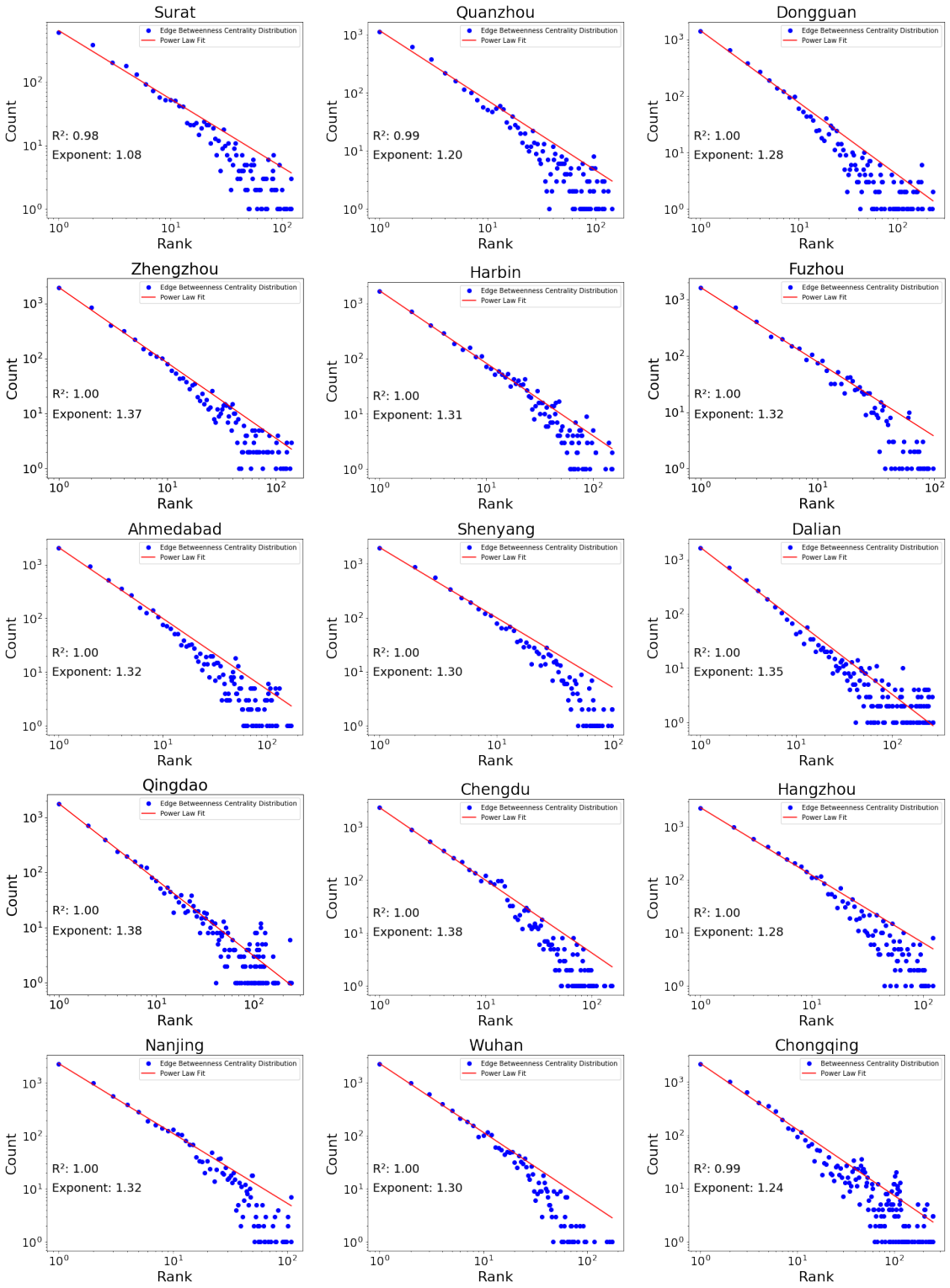


Figure 3.14: Edge betweenness centrality distribution (personal collection 2024).

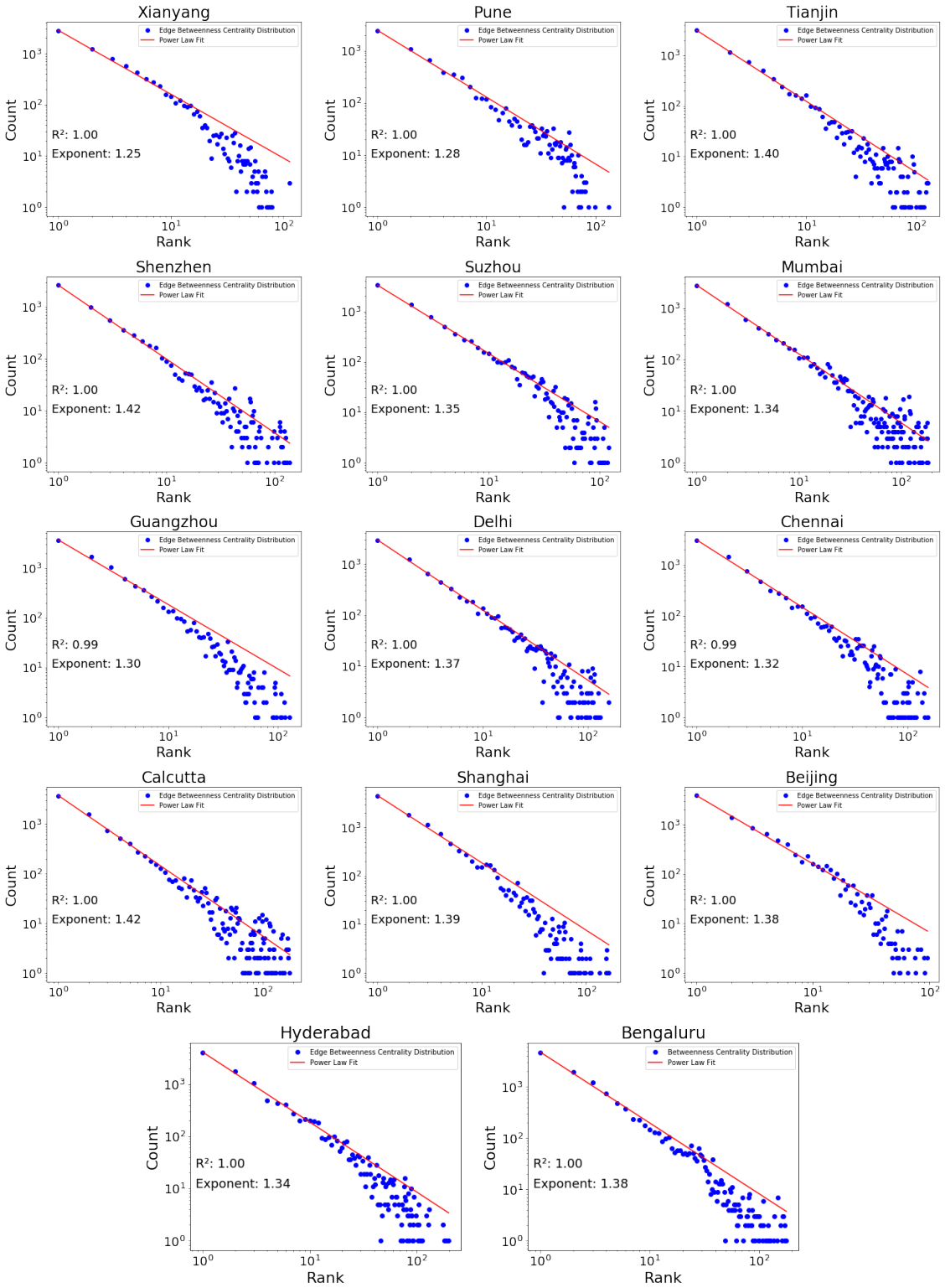
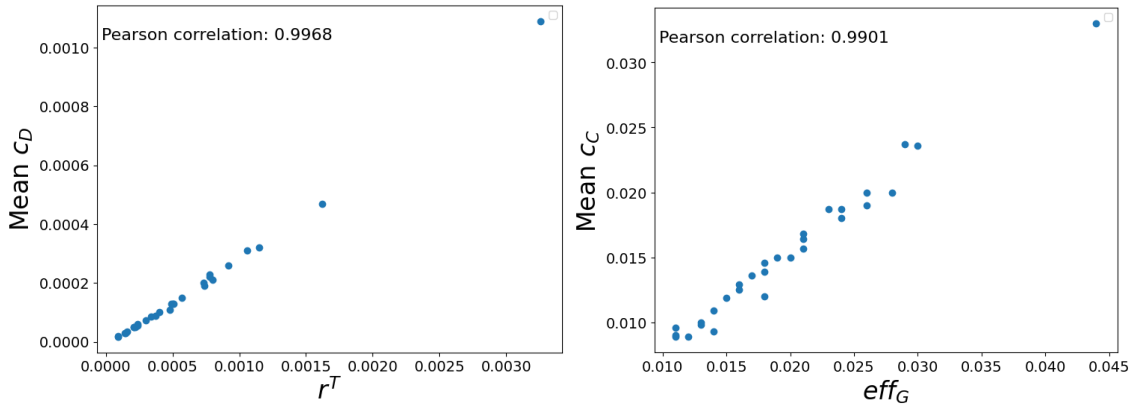


Figure 3.15: Edge betweenness centrality distribution (personal collection 2024) (continued).

A higher closeness centrality value for a road segment implies that it is closer to other nodes in terms of network distance (Lin and Ban 2017; Lan et al. 2022). The closeness distributions, presented in Figure 3.17 and Figure 3.18, offer yet another picture to describe the networks. Broadly, we find that the observed distributions[3] to be akin to a binomial distribution with several cities exhibiting left-skewedness in their distribution. Studying the evolution of Hong Kong road networks between 1976 and 2018, Lan et al. (2022) showed that the closeness distributions over different years can be fitted to a normal distribution with increasing adjusted R^2 as the city grows over the years. Coupled with our observations, we conjecture that the closeness distribution tends to fit normal distribution when the city road network grows denser.

Interestingly, we also found that several Chinese cities also show left-skewed distribution (i.e., Quanzhou, Zhengzhou, Harbin, Shenyang, Qingdao, Hangzhou, Nanjing, Chongqing, Xianyang, and Beijing). The tail extending towards the lower closeness values suggests that in these networks, there are more nodes with longer distances away from the most central nodes. Finally, similar to Lin and Ban (2017), we found that the average closeness size decreases while the size of the road network increases.



(a) r^T and mean c_D (b) eff_G and mean c_C

Figure 3.16: Correlations between macro- and micro-scale metrics.

As in the previous section, we also cross-examined if there are correlations be-

[3]For this, disconnected nodes (with closeness ≈ 0) were ignored.

tween micro-scale and macro-scale metrics. From our analysis (see Figure 3.16), we found that average degree centrality strongly correlated with robustness indicator r^T achieving $\rho(\overline{c_D}, r^T) = 0.9968$ while average closeness centrality positively correlates with efficiency eff_G with $\rho(\overline{c_C}, eff_G) = 0.9901)$.

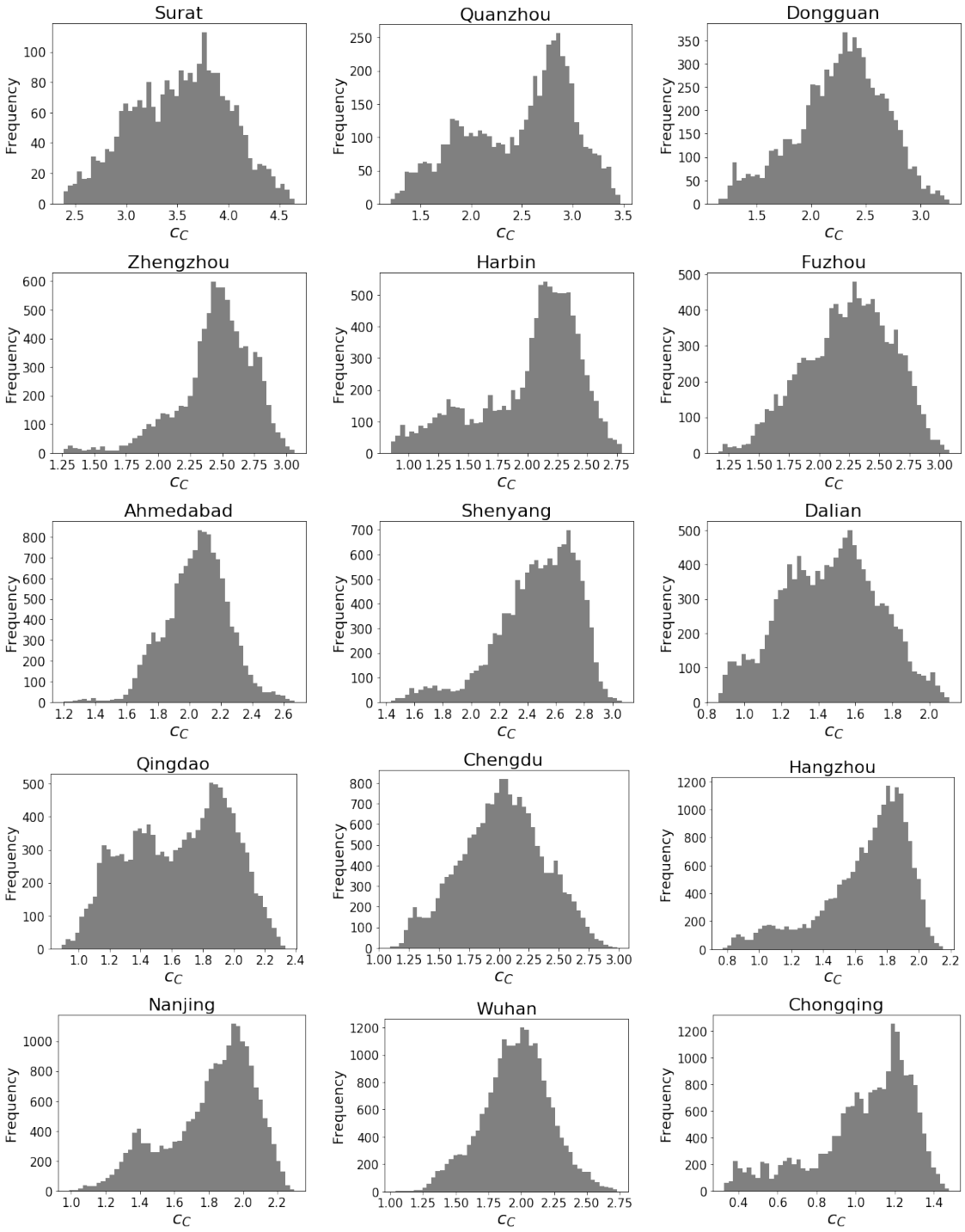


Figure 3.17: Closenness centrality distribution. X-axes have been scaled by a factor of 10^2 for the ease of presentation (personal collection 2024).

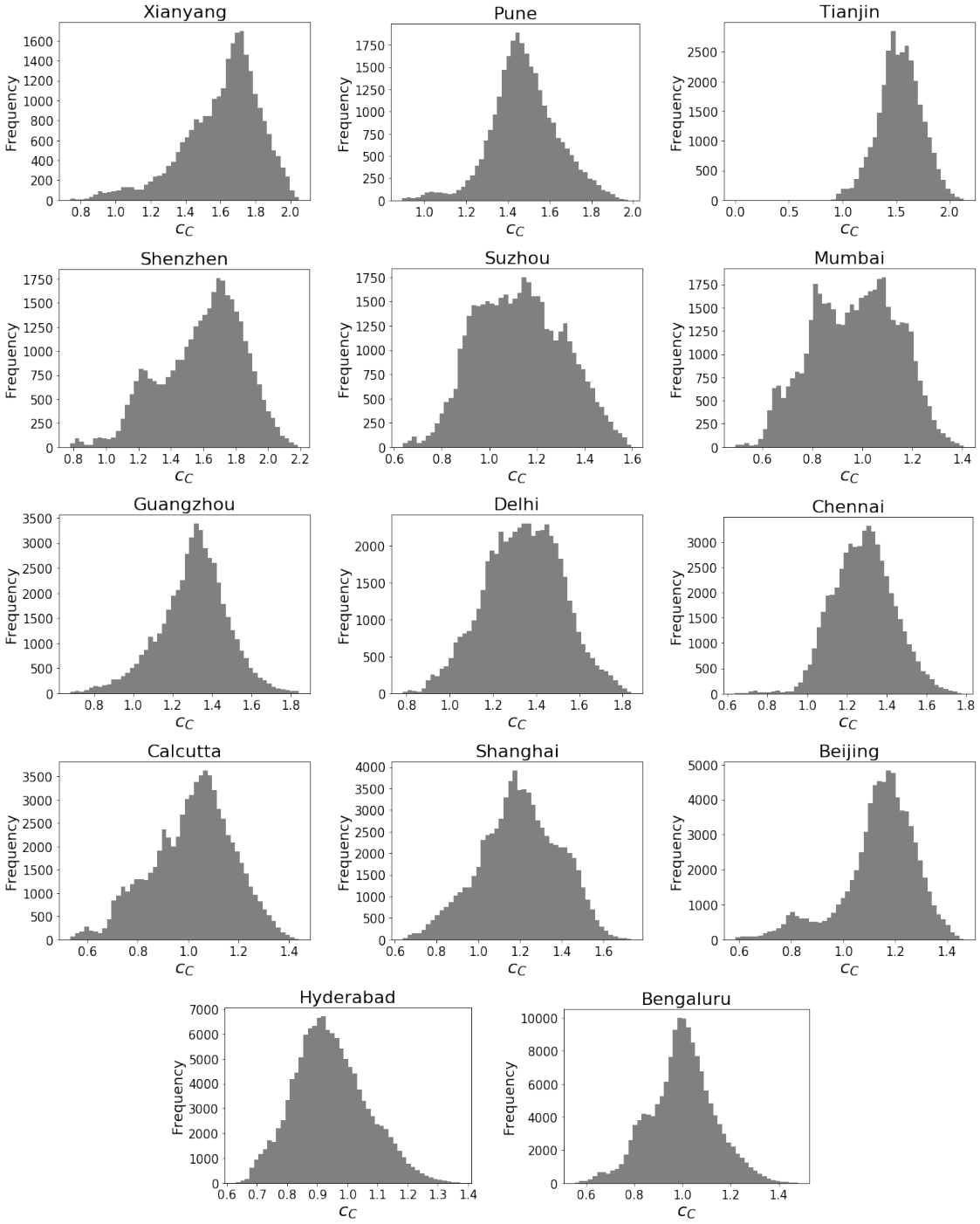


Figure 3.18: Closenness centrality distribution. X-axes have been scaled by a factor of 10^2 for the ease of presentation (personal collection 2024) (continued).

3.6 Chapter Remarks

This chapter presents a topological study of road networks in densely populated cities. Using tools from network science, we first view the road networks as graphs consisting of nodes and links and then approach the analysis in three different scales, namely macro-scale, meso-scale, and micro-scale.

At the macro-scale, we used a range of metrics based on different aspects of the networks (i.e., based on several nodes and links, based on local connectivity, and based on paths). Among these, we find three groups of metrics which are highly positively correlated, (m, $\bar{d}$, κ), (CC_G, τ_G) and (r^T, eff_G), while D and H_G are negatively correlated with other metrics. We note that robust networks (indicated via r^T) are also the more efficient ones (eff_G).

From our analysis at meso-scale, we found that the considered road networks represent a specific class of networks with particularly high modularity which is not often observed in other real-world networks. The high number of communities as well as the low sign of core-periphery structure also suggest that these cities are polycentric with multiple regions having a concentration of road connectivity likely due to high demand to access these regions (e.g., business districts). A future direction on this front would be to consider multiple core-periphery pairs, for instance, using the definition proposed in Kojaku and Masuda (2017).

At the micro-scale (i.e., nodal level), we find the centrality measures considered showing different but consistent distributions. The degree distributions show a lack of degree diversity and highlight the physical spatial constraint in the construction of road infrastructure where most junctions follow the T-shape and there is a very limited number of junctions with high degrees. This then resulted in the distributions not following the power-law which is the opposite to the betweenness distributions which show a strong positive fit to the power-law. We found that the power-law exponent of these cities ranges between 0.98 and 1.39 which are steeper than networks such as the Internet (see Faloutsos et al. 1999). Closeness distributions, on the other hand, show (left-skewed) binomial distribution.

The results and observations from this work offer a multi-scale analysis of the topological properties of road networks in densely populated cities. Our approach offers a set of multifaceted pictures of urban road networks, enabling a more exploratory avenue to understand road networks. Methodologically, this approach is generic and could always be further expanded to include more metrics (e.g., other centrality metrics such as Katz, and PageRank).

Our analysis here should offer insights for relevant stakeholders for understanding, planning, and developing cities. For instance, both road networks in Bologna and central Nantes are monocentric systems (Viana et al. 2013) while in contrast, highly urban road networks studied here exhibit polycentrism. This would inform city authorities/planners whether to decentraliese city development or how to distribute resources across different regions in the city. Finally, our insights (e.g., the distribution shapes of the centrality measures, power-law exponent, and modularity) could also be exploited for modelling and assessing similar classes of networks. The distributions of different centrality measures provide insights into various aspects such as robustness and accessibility which have been used in previous studies such as Lin and Ban (2017) studied the closeness centrality distribution in which different accessibility patterns are found in self-organised cities such as Nanjing and strictly planned cities such as Toronto.

Chapter 4

Robustness of Road Network under Perturbations

In this chapter, we delve into assessing the robustness of urban road networks within densely populated cities. By employing real-world data and conducting a thorough robustness evaluation that involves an iterative node removal process, we closely monitor how the network's structure and integrity withstand the removal of nodes based on their centralities.

This analysis as the previous Chapter 3 focuses on the network's static properties and examines how it responds to disruptions, emphasising its robustness to changes in node configurations. Our investigation encompasses seven distinct node removal strategies, each providing valuable insights into the network's ability to withstand perturbations. Among these strategies, we distinguish between stochastic approaches characterised by random selection and deterministic methods that utilise centrality measures such as degree, betweenness, closeness, Katz, and load centrality.

Through our findings, we unearth the varying impact of perturbations on the network metrics, with random disruption strategies showcasing a relatively lesser detrimental effect compared to targeted disruption strategies based on centrality

measures. Notably, the interplay between different disruption strategies reveals intriguing correlations, with specific combinations demonstrating a balanced disruption sequence. Our analysis underscores the nuanced influence of centrality-based disruption strategies on diverse robustness metrics, shedding light on the intricate dynamics of urban road network robustness.

4.1 Robustness Metrics

As mentioned above in Chapter 2, we follow the definition of robustness as the ability of the network to sustain adequate functionality under network perturbations. In this study, we focus on node removal as the perturbations. To measure robustness, we employ three real-valued metrics, namely the size of the largest connected component (LCC), S_{LCC}, global efficiency, E_{Glob}, and local efficiency, E_{Loc}. We normalise them into range $[0, 1]$ where a higher value indicates a higher functioning network (i.e., 1.0 indicates an unperturbed network implicitly assuming that the original input network has 100% functionality). We compute and track the performance of these metrics over the increasing amount of perturbations introduced to the network under study and gain insights into the behaviour of the network until the networks are fully dismantled.

4.1.1 Size of the largest connected component

The LCC is a fundamental and widely used metric for assessing network robustness. It is defined as the size of the biggest component of a network:

$$S_{LCC} = \max(S_j) \tag{4.1}$$

where S_j is the size (number of nodes) of the j-th component. S_{LCC} is simple to compute and has been used in various works for assessing the global topological connectedness of the network (e.g., Bellingeri et al. 2020; Albert et al. 2000; Duan and Lu 2014; Diop et al. 2022).

4.1.2 Global efficiency

The global efficiency of a network, G, is the average of efficiency between all pairs of nodes whereby the efficiency between nodes i and j is simply the reciprocal of the shortest path length between them. Following this definition, global efficiency of network G, E_{Glob}, can be written as follows (Latora and Marchiori 2001):

$$E_{Glob}(G) = \frac{1}{N(N-1)} \sum_{i \neq j \in V} \frac{1}{H_{i,j}} \tag{4.2}$$

where H_{ij} denotes the shortest path distance between node (in hopcount) v_i and v_j. Essentially, E_{Glob} indicates the effectiveness of traffic exchange within a network. A higher value of E_{Glob} signifies higher efficiency in global communications, as illustrated by Latora and Marchiori (2001). Similar to S_{LCC}, E_{Glob} has been commonly used in assessing network robustness (e.g., Barabasi 2014; Latora and Marchiori 2001; Koulakezian et al. 2015; Manzano et al. 2012).

4.1.3 Local efficiency

While E_{Glob} considers the entire network as a whole, local efficiency instead focuses on the immediate neighborhood of each node. Specifically, the local efficiency of node i is the average efficiency of the local subgraph consisting only of node i's immediate neighbors, but not the node itself. We can then compute the network's local efficiency by taking the average of the local efficiency of all nodes in the network as follows (Latora and Marchiori 2001):

$$E_{Loc}(G) = \frac{1}{N} \sum_{i \in V} E(G_i) \tag{4.3}$$

where

$$E(G_i) = \frac{1}{k_i(k_i-1)} \sum_{l \neq m \in G_i} \frac{1}{H'_{lm}}. \tag{4.4}$$

G_i is the local subgraph of node i. If node i has k_i neighbors, then G_i has k_i

nodes and at most $k_i(k_i - 1)/2$ edges. There, H'_{lm} is the shortest distance between nodes l and m calculated on the graph G_i. As opposed to S_{LCC} and E_{Glob}, to the best of our knowledge, E_{Loc} is not a commonly used metric in the literature. However, we find interesting insights with regard to this metric (cf. Section 4.4). Having introduced the above, we would also like to note that the choice of robustness metrics is dependent on the specific application domain and our study and approach here can naturally be extended to accommodate any other metrics.

4.2 Perturbation Models

A common approach to assess the robustness of a network is by determining how the network responds to different kinds of disruptions as discussed in Chapter 2.

Here, we consider two types of perturbations:

1. Random disruptions corresponding to random events such as accidents (see Chapter 2 Subsection 2.2 for more examples),
2. Targeted disruptions where the intention is to stress test the network by targeting the removal of nodes which are deemed to be the most important in the network to maximise the damage or impact to the network (see Chapter 2 Subsection 2.2 for more examples).

4.2.1 Random disruptions

A road junction may be blocked due to various unplanned reasons (e.g., traffic accidents, fallen trees, landslides). For such events, the sites of occurrence are usually not predictable and they can be modeled as random events.

For this, we consider two types of random node removal strategies.

- Random Point (R_P) – A node is selected at random from all remaining nodes in the network, with each node having an equal probability of being chosen. For this, a node can be removed at any part of the network at each iteration, reflecting well the randomness of events such as accidents.

- Random Area (R_A) – Start by randomly removing a node in the network. In the next iteration, choose to remove a random neighbor of the last removed node. Repeat the process of removing random neighbors iteratively until the desired fraction of node removals is achieved. Such a process reflects a scenario where an area is impacted by events such as flooding or strike actions which usually start at one point in the network and gradually spread from the initial affected location.

4.2.2 Targeted disruptions

It is often important to understand how much disruptions a network can withstand (i.e., worst-case scenario). For such purpose, disruption must be introduced with the aim of maximising damage to the network. To achieve this, we rank the nodes based on their importance in descending order and remove them from the network in that order. To compute the node ranking, we consider different centrality measures, conventionally used in network robustness studies (i.e., Van Mieghem et al. 2010; Kumar and Singh 2020; Trajanovski et al. 2013; Chai et al. 2016).

- Degree centrality (c_D) (Crucitti et al. 2006) – measures the number of direct neighbors each node has. In the context of road networks, it represents the number of road segments (link) meeting at an intersection (node). An intersection with many roads converging towards has a higher degree and has more influence on the local connectivity.

- Betweenness centrality (c_B) (Crucitti et al. 2006) – measures the involvement of a node between all node pairs in the networks (i.e., lies in the shortest path, acting as a bridge between the two nodes). An intersection with high betweenness implies that the location lies on many shortest paths between other nodes and thus, is likely to see a high volume of traffic across this node.

- Closeness centrality (c_C) (Crucitti et al. 2006) – assesses how proximate a node is, on average, to all other nodes in the network. Applied to road networks, a node with high closeness centrality would mean that it can reach

other nodes with fewer hopes.

- Katz centrality (c_K) (Newman 2012) – computes the centrality for a node based on the centrality of its neighbors. It assesses the impact of a node in a network by taking into account both its direct neighbors and all other nodes in the network that connect to the node under consideration through these immediate neighbors. In the realm of road networks, Katz centrality can pinpoint nodes with indirect ties to prominent nodes, shedding light on their secondary level of significance.
- Load centrality (c_L) (Song et al. 2015) – assumes that every node in a network sends an equal amount of a specified commodity to every other node in the network, without considering any capacity limits of edges or nodes. It measures the total amount of flow passing through a node in a network. As opposed to betweenness centrality which ranks node based on its position in the shortest paths between node pairs, load centrality looks at the total flow passing through a node assuming an equal distribution.

The definitions of the above-mentioned centrality measures (i.e., c_D, c_B, and c_C) were introduced previously in Section 3.3 and the rest (i.e., c_K and $c_D L$) are given in the Table 4.1.

Table 4.1: Centrality measures used to rank nodes for targeted disruptions

`Katz`	$c_K(v_i) = \beta(I - \alpha A)^{-1}$
`Load`	$c_L(v_i) = \sum_{v_j, v_k \in V} \frac{1}{\sigma(v_j, v_k \mid v_i)}$
$\sigma(v_j, v_k \mid v_i)$ is the number of those paths between v_j and v_k passing through node v_i; α is a constant (damping factor), usually $\alpha < 1/\lambda$ where λ is the largest eigenvalue of adjacency matrix. When $\alpha \geq \lambda$, the centrality tends to diverge; β is a bias constant (exogenous vector) used to avoid the zero centrality values; I is the identity matrix.	

4.3 Network Perturbation Analysis

Our analysis involves the introduction of perturbations to the network under study. For this, we sequentially remove an increasing fraction of nodes from the network following the perturbation models described in Section 4.2. For random disruptions (i.e., R_P and R_A), since they are stochastic, we repeat the experiment for each network 100 times and present the mean value along with their 95% confidence interval using error bars.

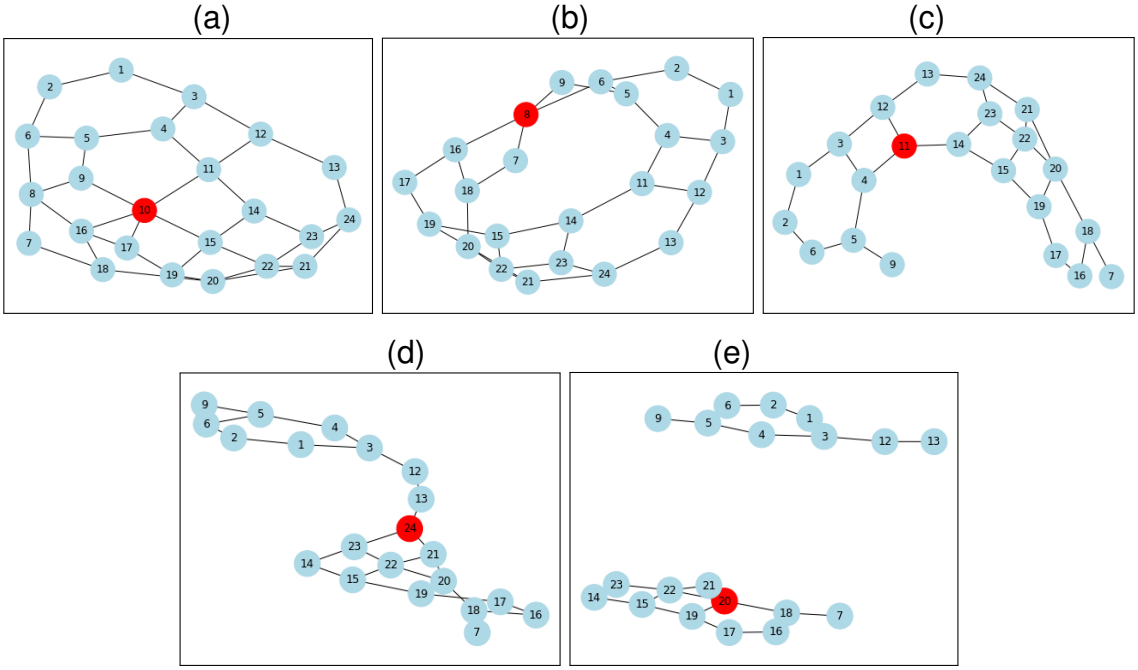


Figure 4.1: Illustration of node removal process based on c_B for a sample 24-node road network where the node with the highest c_B for each LCC is indicated in red (personal collection 2024).

Figure 4.1 provides an illustration of the removal process following node ranking based on betweenness centrality. For each step, the node with the highest c_B (node in red color) is removed. Note that after four nodes were removed, the network disconnected into two components (see Figure 4.1(e)). When this happens, following the literature (e.g., Trajanovski et al. 2013), we continue to consider only the new LCC and ignore the small component(s).

Algorithm 1: Node disruption Algorithm

Input : Network $G = (V, E)$, X=$\{R_P, R_A, c_D, c_B, c_C, c_K, c_L\}$

Output: S_{LCC}, E_{Glob}, E_{Loc}

Initialization:
Compute and record S_{LCC}, E_{Glob} and E_{Loc} of G;
Let LCC = G
if $X == R_P$ **then**
    while $S_{LCC} \neq 0$ **do**
        Node = getRandomNode(LCC);
        remove Node from LCC;
        record new S_{LCC}, E_{Glob}, E_{Loc} of G;
    end
end
else if $X == R_A$ **then**
    Node = getRandomNode(LCC);
    while $S_{LCC} \neq 0$ **do**
        remove Node from LCC;
        record new S_{LCC}, E_{Glob}, E_{Loc} of LCC;
        Node = getNeighbor(Node);
    end
end
else
    nodeList = rankNodes(G,X) //put nodes in LCC in decreasing order based on X;
    index=0;
    while $S_{LCC} \neq 0$ **do**
        remove nodeList(index);
        index=getNextNode(nodeList);
        record new S_{LCC}, E_{Glob}, E_{Loc} of LCC;
    end
end

We track the gradual degradation of the robustness metrics introduced in Section 4.1 (i.e., the resulting S_{LCC}, E_{Glob}, and E_{Loc}) after each node removal and recompute the LCC. We consider this new LCC for the next removal. Algorithm 1 presents the pseudocode for our node removal process.

4.4 Robustness Assessment

4.4.1 Dataset

In the context of this research, a comprehensive selection process was undertaken to identify ten distinct road networks from the dataset introduced in detail in Section 3.4. These road networks include representations from two cities in India, specifically Surat and Ahmedabad, as well as eight cities in China, namely Quanzhou, Dongguan, Zhengzhou, Harbin, Fuzhou, Shenyang, Dalian, and Qingdao.

4.4.2 Correlation and similarity of different targeted disruption strategies

Before we present our robustness assessment, we first offer some insights into the different targeted disruptions (i.e., the node rankings based on different centrality measures). For this, we first investigate the extent to which different centrality measures result in the removal of similar nodes. We compute the Spearman coefficient as a full-rank correlation proxy (see Figure 4.2). We see a strong correlation between c_B and c_L and between c_D and c_K, implying that one from each pair could be redundant in future analysis as they almost provide the same node removal sequences. At the other end of the spectrum, we find c_D and c_C to have the lowest correlation while others fall between these two extremes.

We further compute the pairwise similarity of the centrality measures to understand how similar the rankings are when an increasing fraction of nodes are removed. To achieve this, we consider the following process. For two node rankings, $C_a = [c_{a_1}, c_{a_2}, \ldots, c_{a_N}]$ and $C_b = [c_{b_1}, c_{b_2}, \ldots, c_{b_N}]$, $U_{C_a,C_b}(k)$ is the percentage

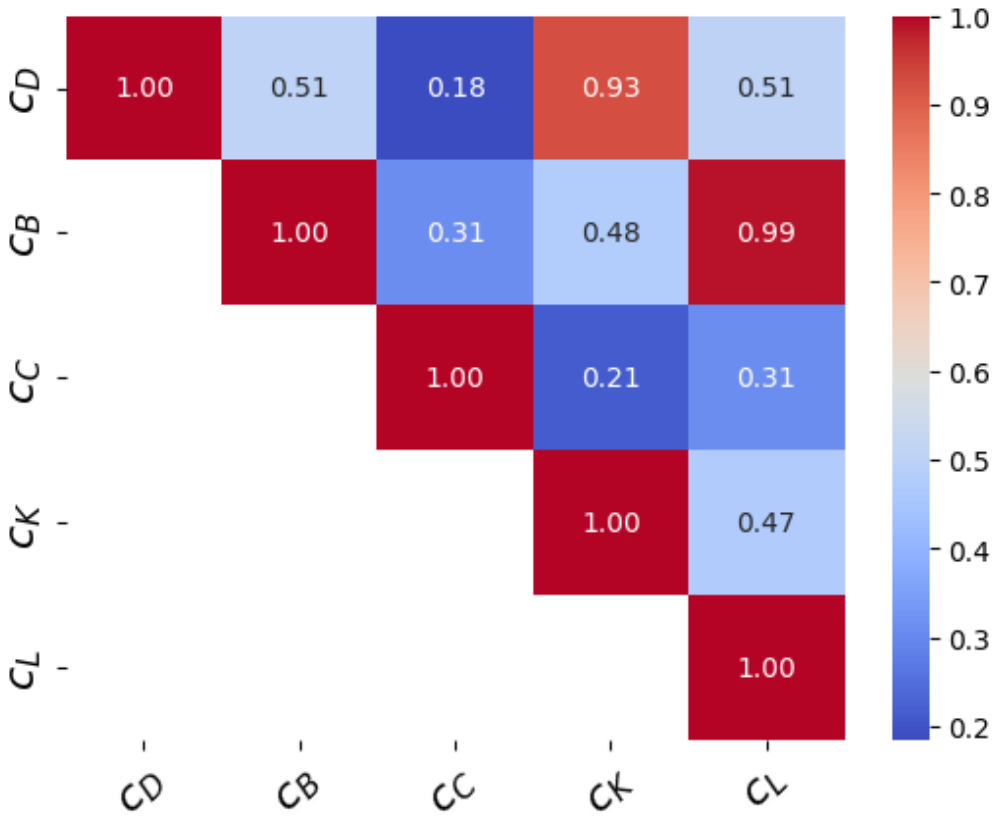


Figure 4.2: Spearman correlation heatmap of the centralities (the targeted disruptions). The heatmap represents the average of all ten networks.

of nodes in $\{c_{a_1}, c_{a_2}, \ldots, c_{a_{\lfloor kN \rfloor}}\}$ that also appear in $\{c_{b_1}, c_{b_2}, \ldots, c_{b_{\lfloor kN \rfloor}}\}$. In this way, when $k = 100\%$, complete overlap is achieved, resulting in $U_{C_a,C_b}(100\%)$=1. Essentially, $U_{C_a,C_b}(k)$ indicates the proportion of shared nodes from the top $k\%$ of nodes in the rankings C_a and C_b. The results of $U_{C_a,C_b}(k)$ for studied road networks are given in Figure 4.3.

Five different centrality measures give us ten possible combinations of centrality pairs. From the figure, we observe that $U_{Betweenness,Load}(k)$ generally has the highest overlap across increasing k. It is followed by $U_{Degree,Katz}(k)$. The remaining eight pairs perform closely with $U_{Closeness,Katz}(k)$ and $U_{Degree,Closeness}(k)$ showing the lowest values for all networks. Since $U_{C_a,C_b}(k)$ is low when C_a and C_b have few overlap (i.e., C_a and C_b consider node importance differently), then pairs of centrality measures with low $U_{C_a,C_b}(k)$ have distinct impacts to the road network. Considering this, with $U_{Closeness,Katz}(k)$ and $U_{Degree,Closeness}(k)$ showing

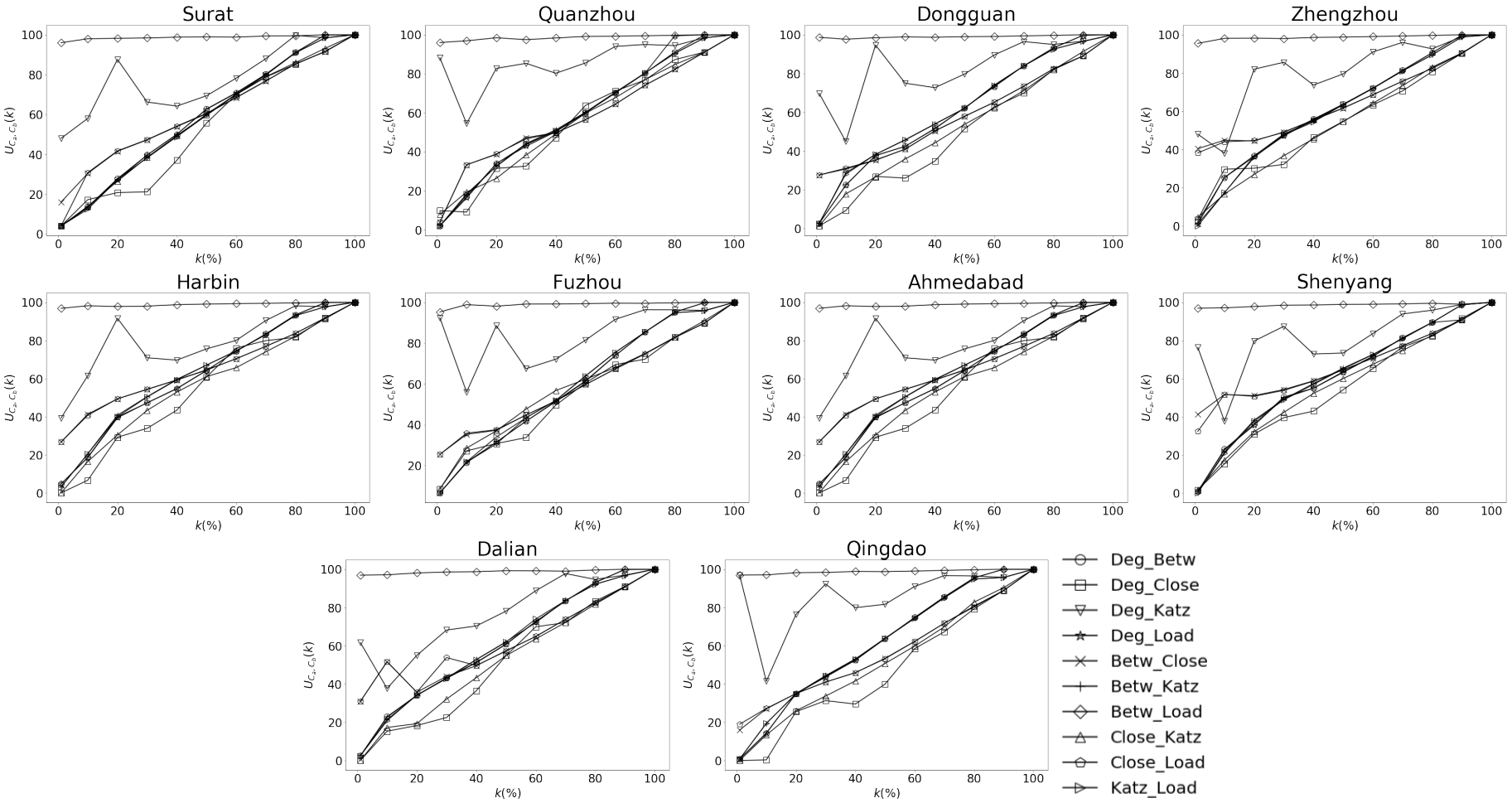


Figure 4.3: Similarities of centrality rankings for road networks. Each plot shows the overlap of nodes (y-axis in %) from the first nodes (x-axis) ranked according to centrality ranking C_a and the first nodes ranked according to centrality ranking C_b for a given network.

relatively few overlaps between their node rankings, they should be considered to gain insights into the impact to the network following different node disruption sequences.

4.4.3 Largest connected component

We begin our assessment focusing on LCC and present in Figure 4.4 the evolution of S_{LCC} in our experiments. Not surprisingly, as we increase the number of perturbations, S_{LCC} decreases monotonically for all networks. However, in general, random perturbations (i.e., R_P and R_A) are less effective in disrupting the network (shown by the slower degradation of S_{LCC}) compared to targeted perturbations based on centralities.

Between R_P and R_A, our results suggest disruptions spreading around a neighbourhood region in the network, R_A, incur a higher detrimental impact than disruptions occurring at random locations in the network. This is likely due to the fact that R_A removes nodes focusing in one area and thus manages to disconnect the network into components more effectively.

Proceeding to consider the targeted disruptions, we first observe that, for all networks, c_K is the least effective in disintegrating the network. This implies that for road networks, a node directly connected to or near important major junctions does not necessarily inflate the importance of the node of interest and c_K is a poor choice to degrade S_{LCC}.

This is followed by c_D. While many real-world networks exhibit scale-free properties with power-law degree distributions (Barabasi 2013), we found that urban road networks do not possess such degree distribution. Rather, since they are spatial networks, they have a small deviation of average degree (between 2.7 to 2.95 with an average standard deviation of 0.94 and average variance of 0.89). This observation has also been found by previous works (e.g., Reza et al. 2022; Lee and Jung 2018; Akbarzadeh et al. 2018) that urban road networks do not exhibit large variation in degrees due to planar constraints (Lämmer et al. 2006; Viana et al. 2013).

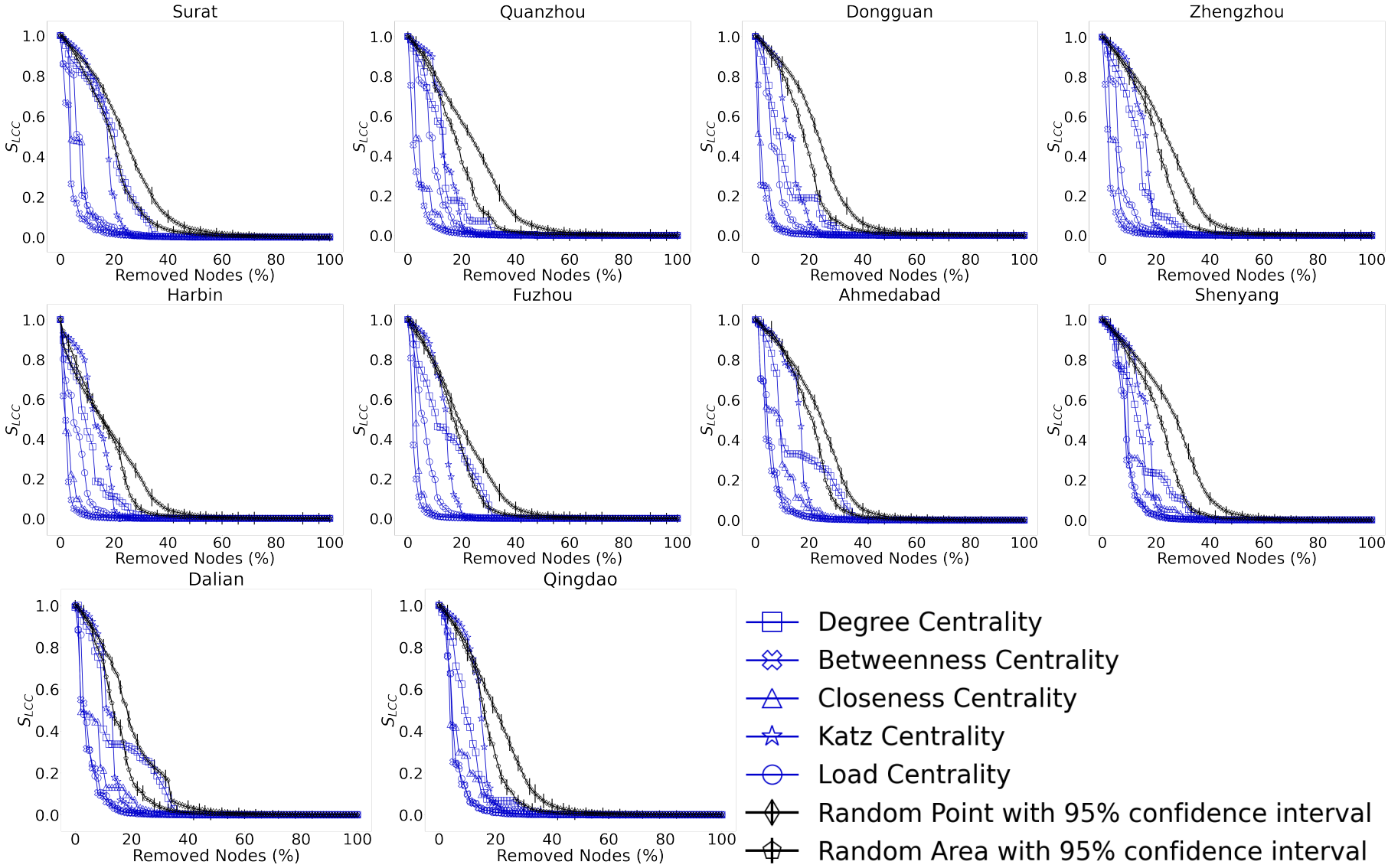


Figure 4.4: The evolution of S_{LCC} curve (normalised) for different disruption strategies in the ten road networks.

In our case, we see that the road networks have a majority of junctions (nodes) inter-connecting two or three roads. Hence, with many nodes having similar degrees, the c_D disruption strategy ultimately does not differentiate most of the nodes. Moreover, we observe clear knee points for the S_{LCC} curves in Figure 4.4, indicating removal of some nodes with similar degrees has a small impact on S_{LCC}.

At the other end of the spectrum, we found that disruption based c_B to be the most effective and this is consistent across all networks. This aligns with previous research in Albert et al. (2000) where it was also found that node removal based on c_B leads to the worst-case scenario for the robustness of complex networks. In the context of transportation, Duan and Lu (2014) found that c_B-based disruption strategy is the most harmful. In Vaca-Ramírez (2019), only ≈10% of nodes removed based on c_B ranking is needed to incur significant deterioration of the road network of Quito city.

Meanwhile, the effectiveness of disruption strategies, c_C and c_L, vary between different networks but overall, c_C appears to be more disruptive for smaller networks while c_L is more effective for bigger ones. However, the differences are marginal. Based on the above discussion, we could broadly summarise the effectiveness of the different disruption strategies in the following order:

$c_B \succ c_C \approx c_L \succ c_D \succ c_K \succ R_A \succ R_P$. [1]

In Figure 4.5, we present the percentage of nodes needed to be removed to achieve 25%, 50% and 90% reduction of S_{LCC} for the considered road networks. From this figure, we can make several further observations. First of all, we can see that, overall, Shenyang appears to have the most robust road network when most disruption strategies are less effective on it than other networks. Delving further into this, we found that Shenyang has the highest degree diversity among all the cities considered here, suggesting high degree diversity offers better robustness in terms of S_{LCC}.

[1] To simplify discussion and presentation, we use $X \approx Y$, $X \succ Y$, and $X \prec Y$ to indicate that X inflict similar, higher, and lower degradation to the network than Y respectively.

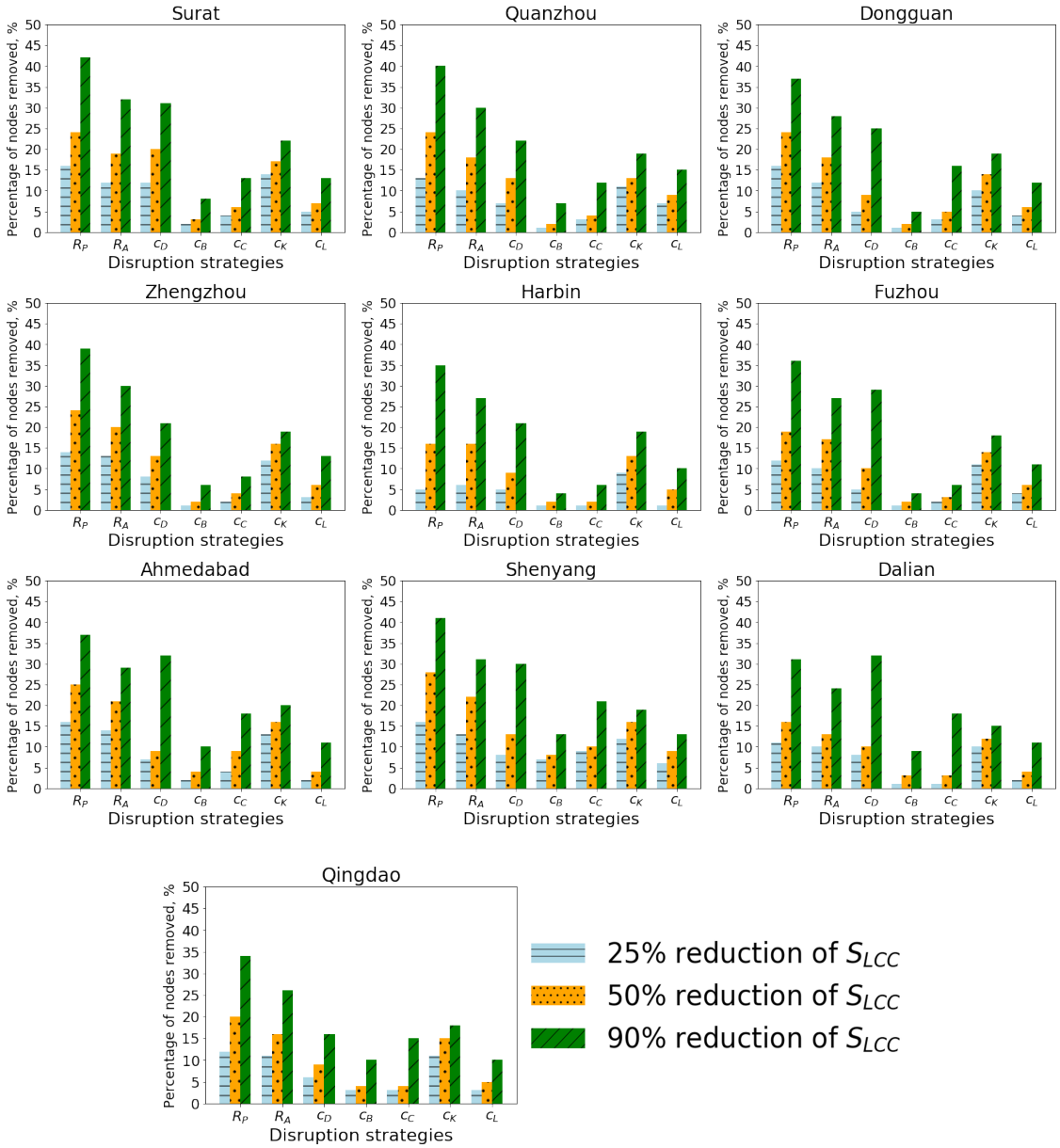


Figure 4.5: Reduction of S_{LCC} based on types of perturbation strategies to achieve 25%, 50%, and 90% decrease.

Second, while c_D-based disruption is generally more effective than random removal strategies, (i.e., R_A and R_P), there are cities (i.e., Fuzhou, Ahmedabad, and Dalian) where a higher number of nodes are needed to be removed to achieve 90% S_{LCC} reduction even though much lower number of nodes are needed to achieve 25% and 50% S_{LCC} reduction. This indicates that disruption based on node degree has smaller impacts when the S_{LCC} is small.

Third, c_K disruption strategy seems to need a similar number of nodes removed to achieve the three 25%, 50% and 90% S_{LCC} reduction; forming a more linear relationship between node removal and decrease of S_{LCC} at the end of the node removal process compared to other strategies. Finally, the least effective R_A require ≈30-40% removal to achieve 90% S_{LCC} decrease.

4.4.4 Global efficiency curve

We present in Figure 4.6 the evolution of E_{Glob} with increasing perturbations for all the cities. Similar to S_{LCC} in Figure 4.4, we can still set apart random and targeted disruption strategies where random ones are less effective in decreasing the E_{Glob} with $R_A \succ R_P$. Broadly, random node disruption leads to a more gradual deterioration until the network eventually collapses when $\approx$ 40% of nodes are removed.

Our findings align with other studies Trajanovski et al. (2013), where similar percentage of nodes were removed to cause a collapse in the complex networks (i.e., industrial networks (Alrumaih and Alenazi 2023), metro networks (Zhao et al. 2018), and power grid (Trajanovski et al. 2013)). However, we observe an exception for Qingdao when R_A is actually more effective than c_D.

In fact, compared to S_{LCC}, the curves for E_{Glob} are closer, indicating the smaller differences in impact for the different strategies. Disruption based on c_B and c_L appear to be the most effective ones with both performing similarly ($c_B \approx c_L$). The rest are close and dependent on the city with c_D marginally worse than others.

We present in Figure 4.7 the percentage of nodes needed to be removed to reduce E_{Glob} by 25%, 50%, and 90%. From the figures, Surat and Quanzhou are

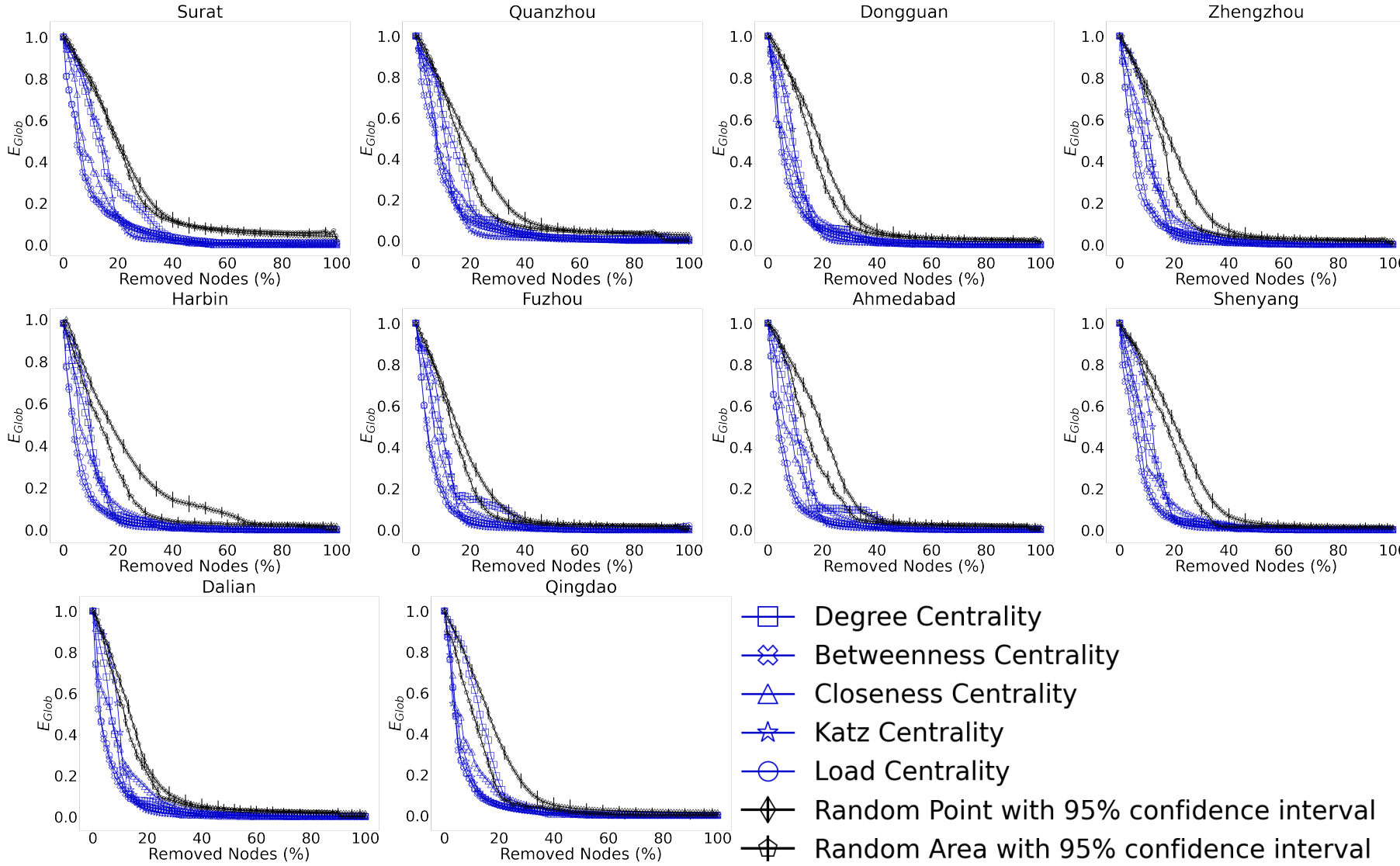


Figure 4.6: The evolution of E_{Glob} curve (normalised) for different disruption strategies in the ten road networks.

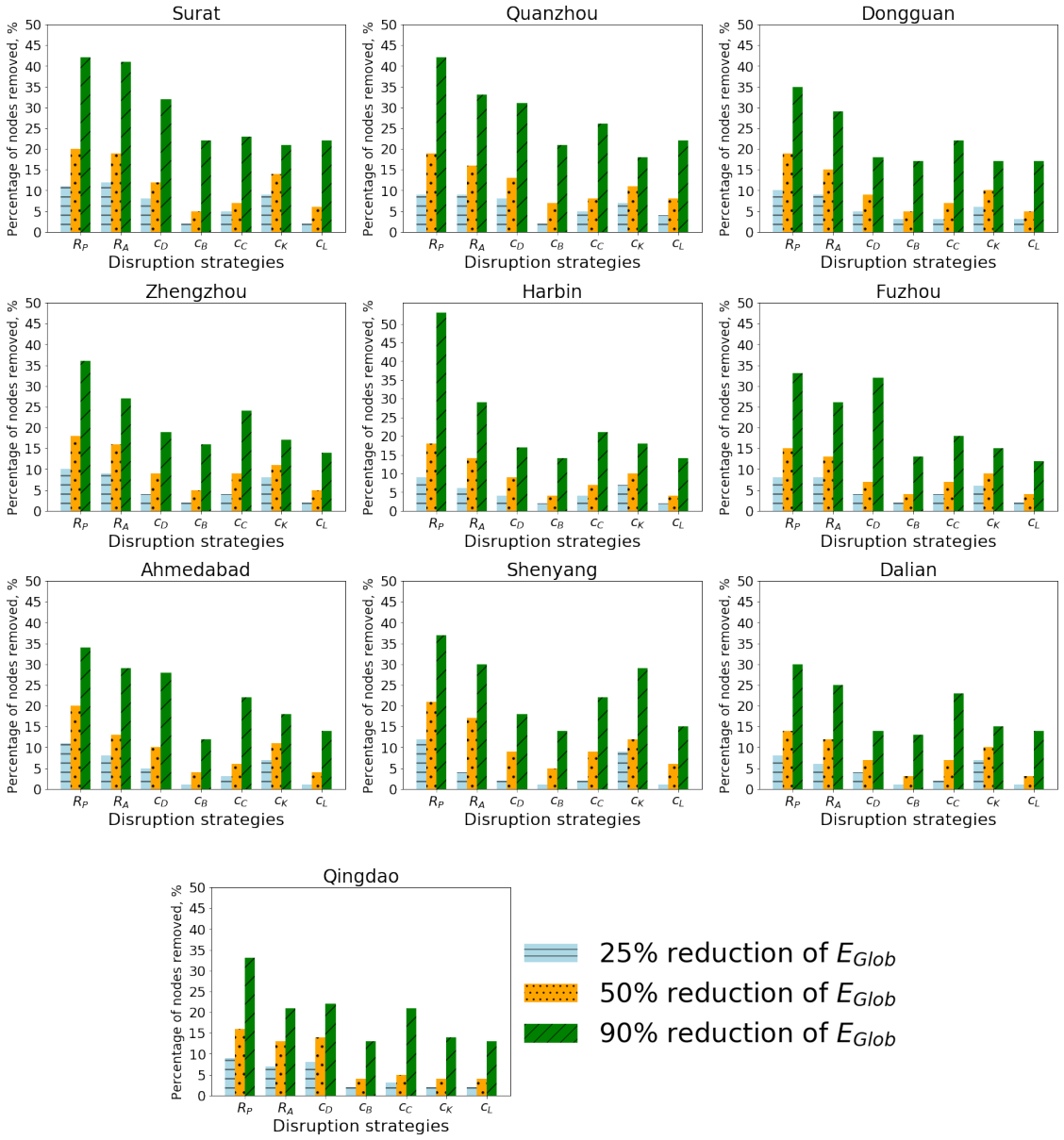
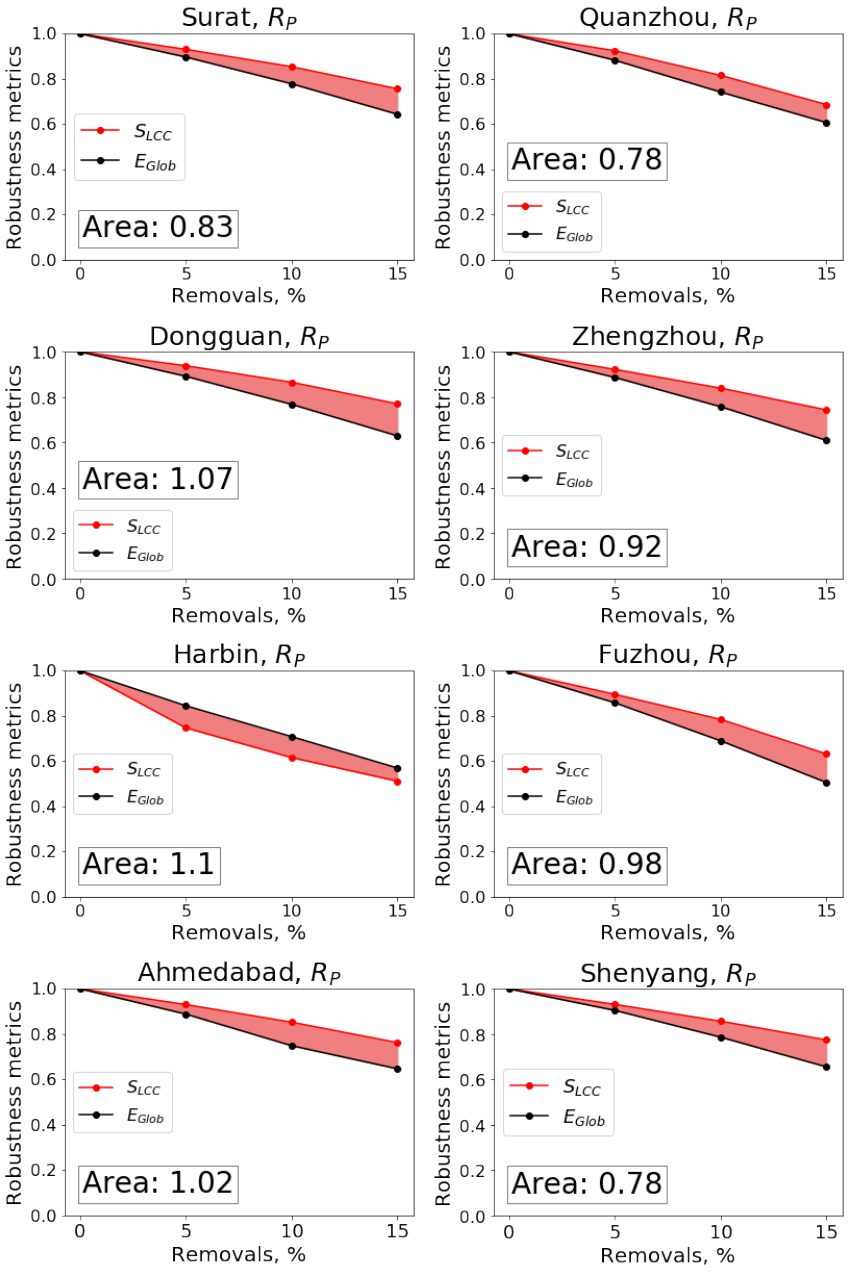


Figure 4.7: Reduction of E_{Glob} based on types of perturbation strategies to achieve 25%, 50%, and 90% decrease.

generally the most robust against random failure R_P and R_A for 25%, 50%, and against targeted disruption strategies. However, for R_P at 90% reduction of E_{Glob}, Harbin is the most robust compared to all studied road networks. Overall, Dalian appears to be the least robust to the R_P and R_A based on E_{Glob}.

We now focus on the initial phase of disruption where only small perturbations are introduced as this is the most important in assessing how a network may maintain its function. A fast deterioration at the initial phase would indicate that

the network's function can be severely disrupted with minimal perturbations. For this purpose, we look at the robustness metrics discussed so far. We present the results in Figure 4.8 - Figure 4.14. In these figures, we compare the difference between the decrease of S_{LCC} and E_{Glob} at 5%, 10% and 15% node removal. The shaded area indicates the difference between the two metrics.



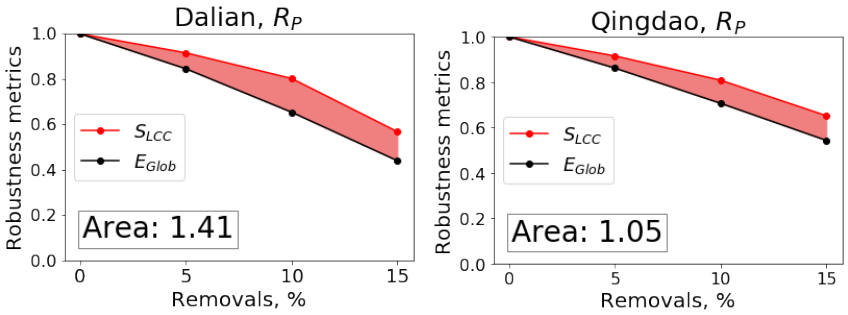
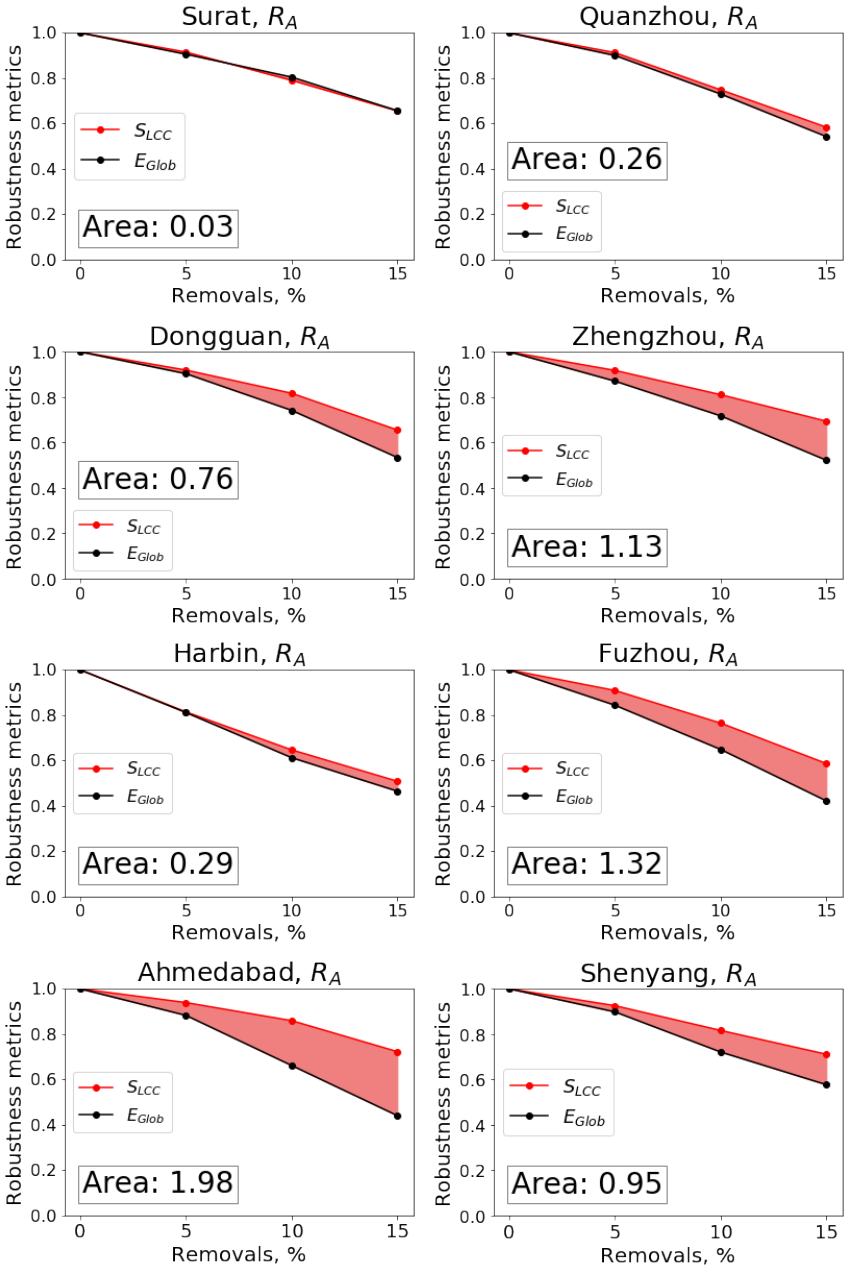


Figure 4.8: Comparing the decrease of S_{LCC} and E_{Glob} with 5%, 10%, 15% of nodes removed with R_P.



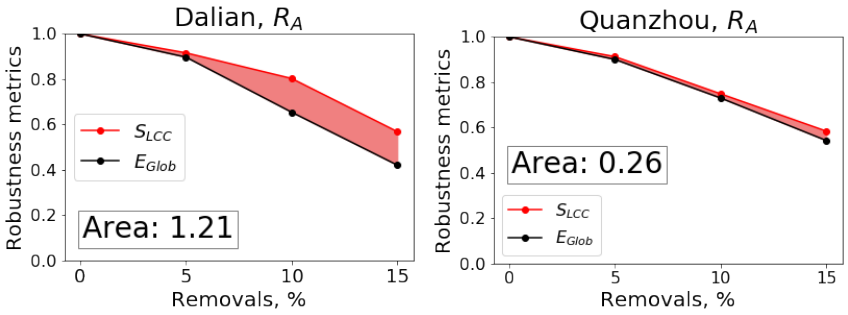
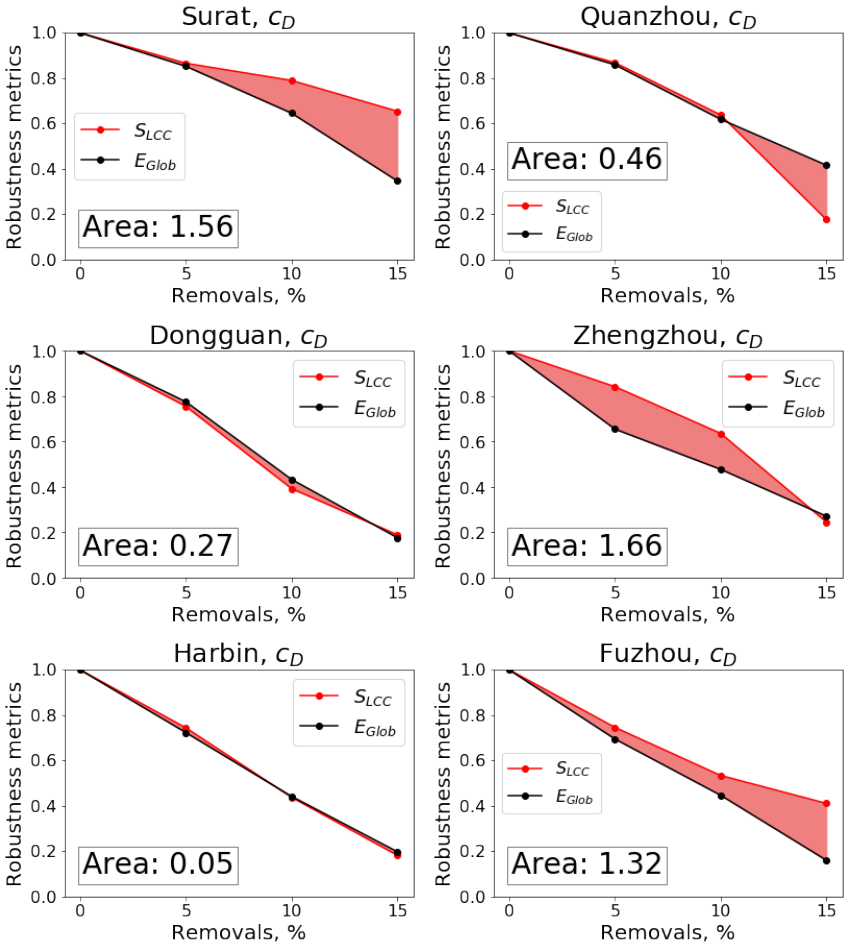


Figure 4.9: Comparing the decrease of S_{LCC} and E_{Glob} with 5%, 10%, 15% of nodes removed with R_A.

Figures Figure 4.8 - Figure 4.9 compare the difference between the decrease of S_{LCC} and E_{Glob} at 5%, 10% and 15% node removal for R_P and R_A, we see increasing differences between S_{LCC} and E_{Glob} for R_P and R_A when N is increasing.



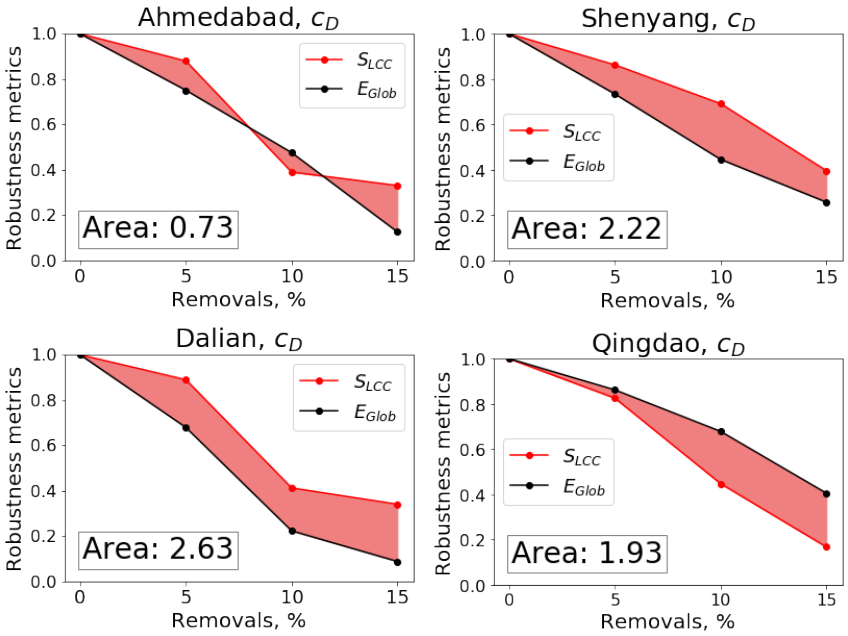
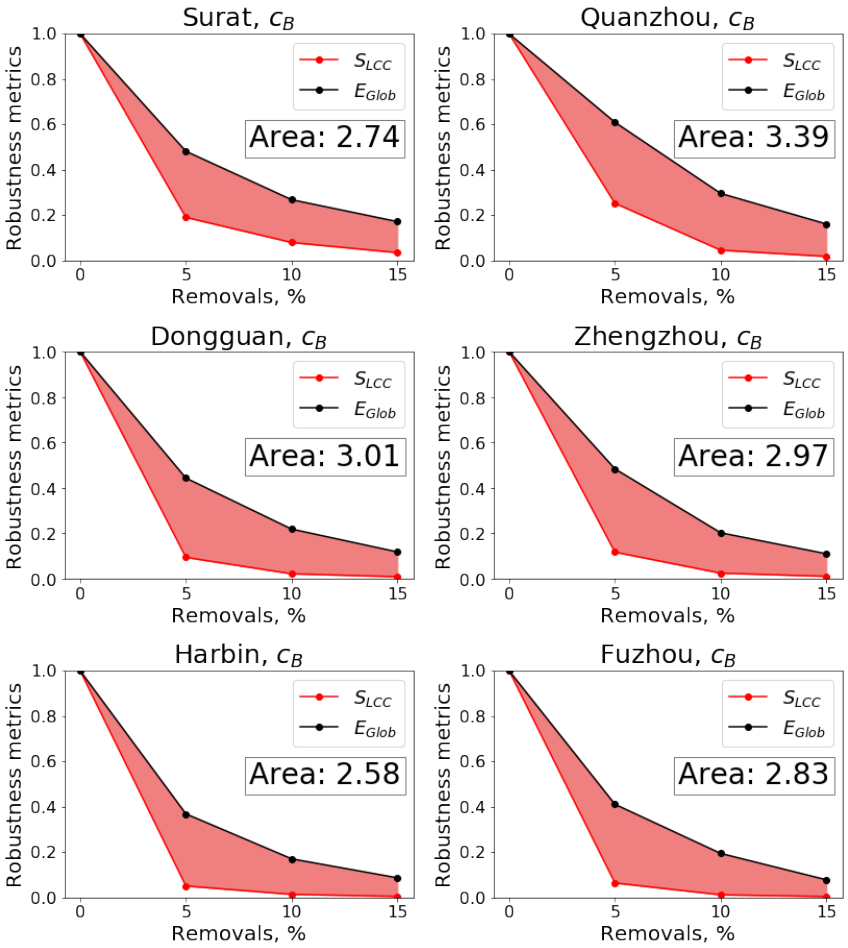


Figure 4.10: Comparing the decrease of S_{LCC} and E_{Glob} with 5%, 10%, 15% of nodes removed with c_D.



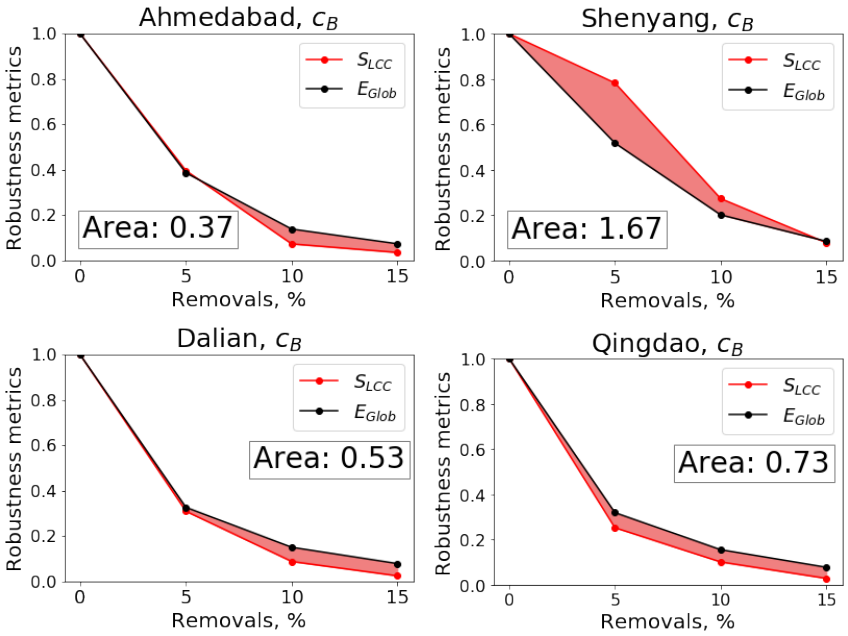
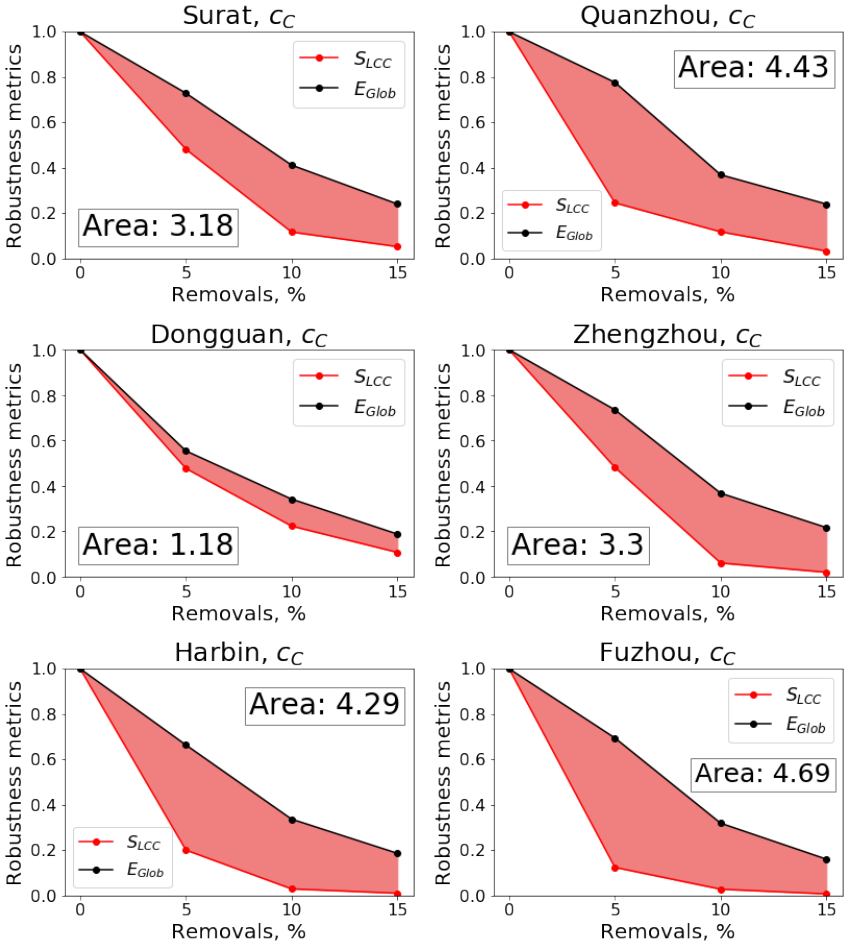


Figure 4.11: Comparing the decrease of S_{LCC} and E_{Glob} with 5%, 10%, 15% of nodes removed with c_B.



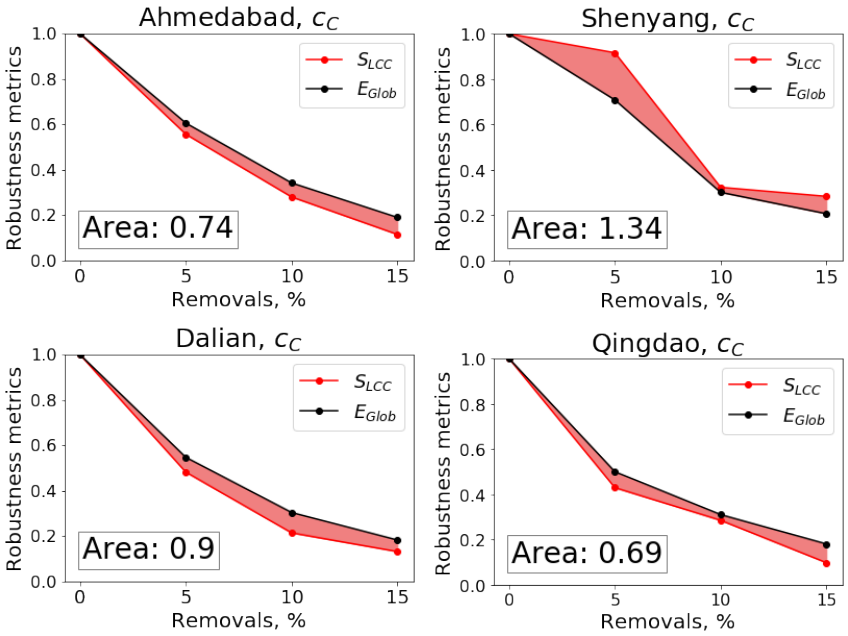


Figure 4.12: Comparing the decrease of S_{LCC} and E_{Glob} with 5%, 10%, 15% of nodes removed with c_C.

From Figure 4.10, we observe increasing differences between S_{LCC} and E_{Glob} for c_D when N is increasing. However, for Figure 4.11 and Figure 4.12 the converse is true for c_B and c_C where we see bigger differences for smaller networks and vice versa. For instance, we note that 15% removal of nodes with the highest c_B or c_C already incur almost 100% of S_{LCC} for smaller networks such as Surat and Quanzhou though E_{Glob} appears to be affected less.

Finally, the differences observed for c_K (in Figure 4.13) and c_L (in Figure 4.14) remain relatively stable across the 10 studied road networks.

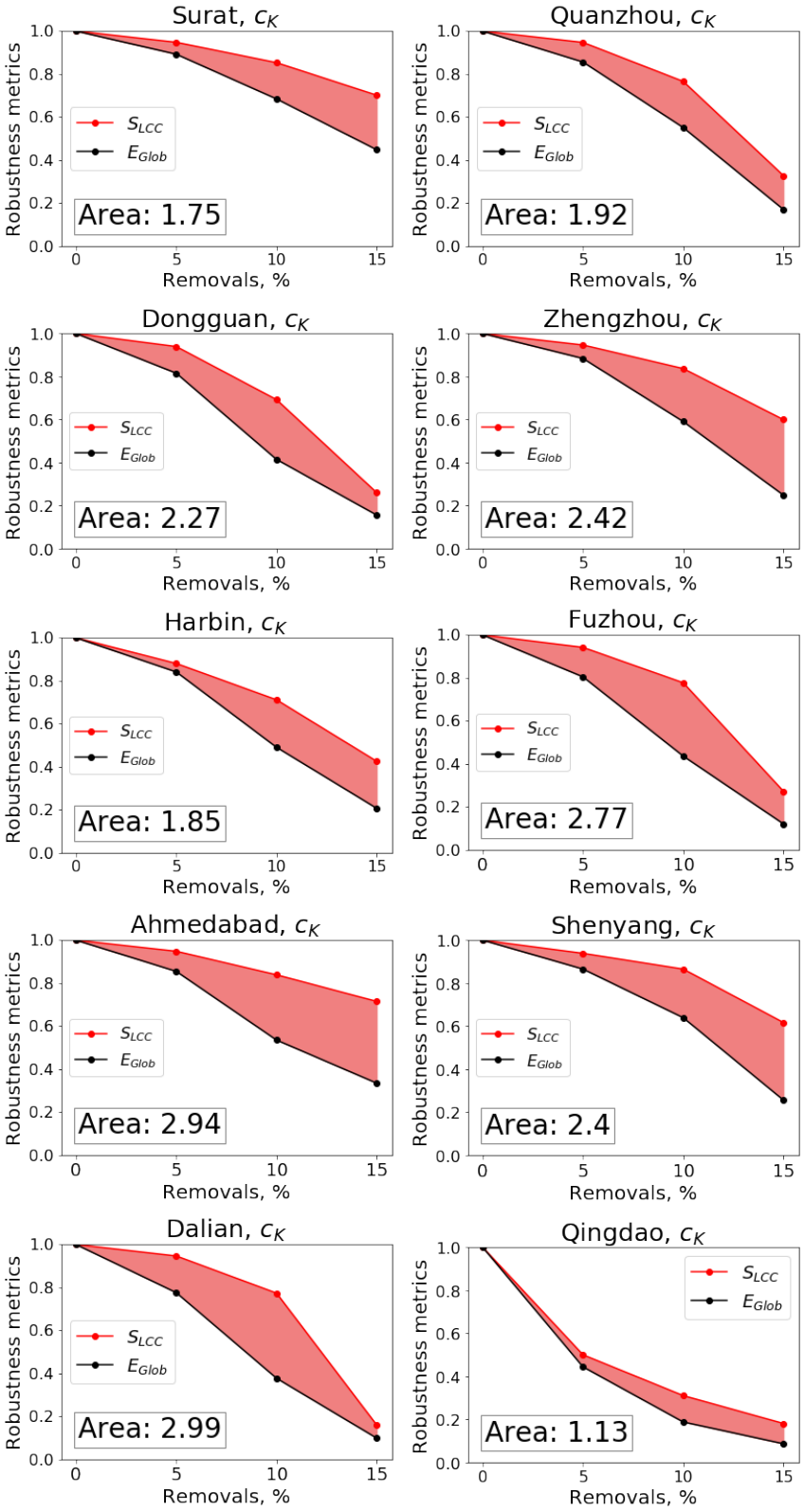


Figure 4.13: Comparing the decrease of S_{LCC} and E_{Glob} with 5%, 10%, 15% of nodes removed with c_K.

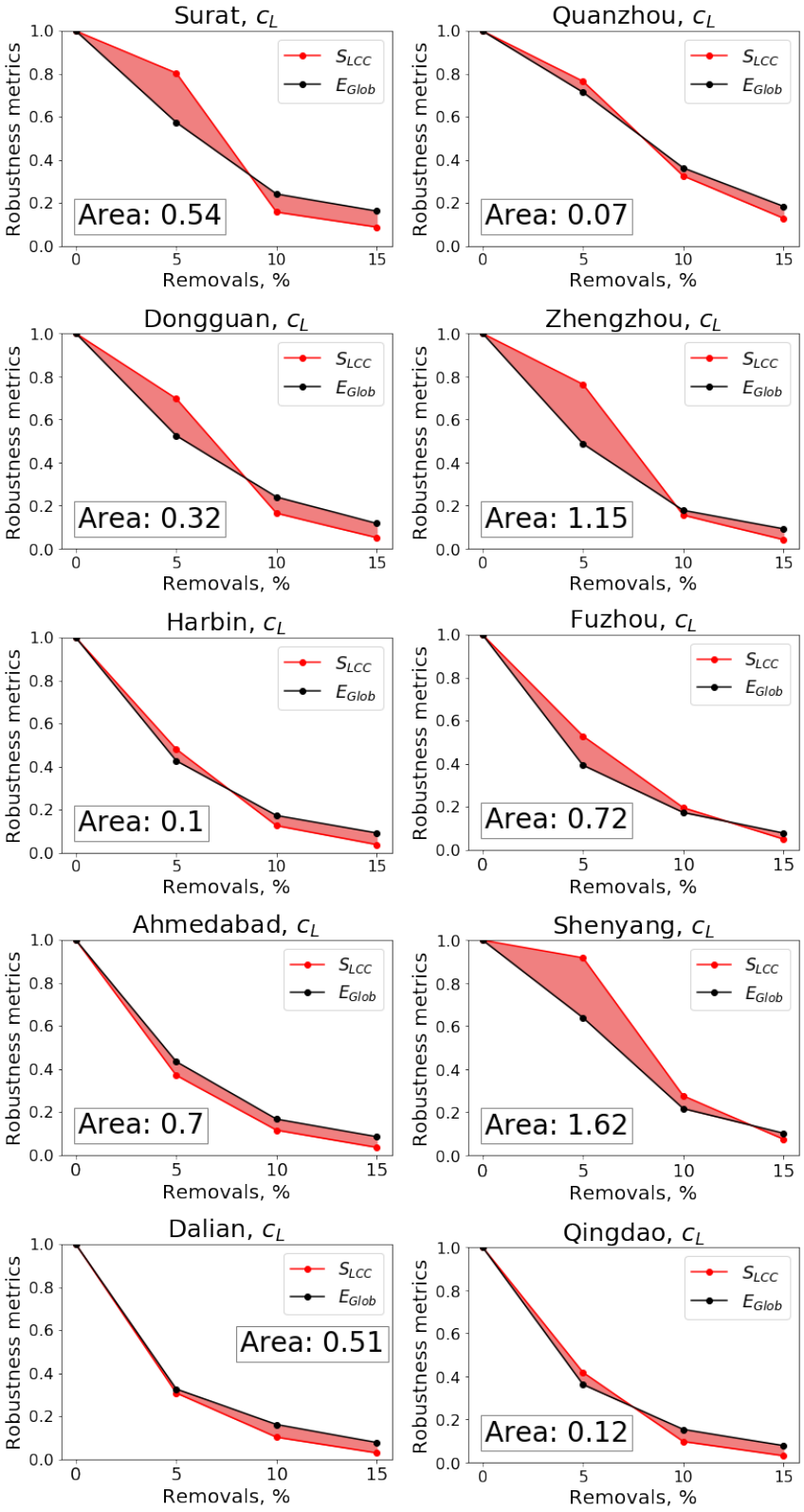


Figure 4.14: Comparing the decrease of S_{LCC} and E_{Glob} with 5%, 10%, 15% of nodes removed with c_L.

4.4.5 Local efficiency curve

In this section, we focus on a less explored metric, i.e., local efficiency E_{Loc}. Instead of viewing the efficiency of the entire network as a whole, E_{Loc} focuses on the efficiency within the immediate neighborhood of each node. This is relevant to local traffic as many car trips in big cities have been found to be short or within small locality (i.e., a third of car trips in London are shorter than 1 km (Transport for London 2012); in 2019, 17% of UK car journeys were between one and two miles (Carlton 2023); nearly 25% of car trips were shorter than five minutes in Sydney (Sugiyama et al. 2012); less than 5 km car trips made up more over 40% of all car trips in 2010 in Beijing (Ming et al. 2014)). There are various initiatives to discourage short car trips to reduce environmental impacts (e.g., replacement with micromobility modes (Fan and Harper 2022; Transport Scotland 2022; Lang World Economic Forum 2022)).

Figure 4.15 shows the evolution of the average E_{Loc} when the network is perturbed based on the seven different perturbation strategies. From the results, we can clearly see three groups of strategies. The first group consists of the random perturbation strategies, R_P and R_A, showing smoother gradual degradation of E_{Loc}. In the case of R_A, it is relatively steady until $\approx 20 - 30\%$ of nodes are removed, followed by a rapid decline.

The most impactful disruption strategies, c_K and c_D, forms the second group where we see steeper decline in E_{Loc}. This is opposite to the S_{LCC} and E_{Glob} metrics where c_K and c_D are the less disruptive ones. Networks collapse when $\approx 30\%$ of nodes are removed based on c_K and $\approx 40 - 45\%$ nodes based on c_D. From these observations, we can deduce that c_K and c_D offers strong disruption to local region but not the overall network where other centrality measures (e.g., c_B) can incur greater damage to the network as a whole.

The third group consisting of three removal strategies based on c_B, c_L, and c_C is perhaps the most interesting one. In contrast with previous metrics, we observe an initial *increase* of E_{Loc} for these removal strategies, implying an improvement of local efficiency after node removals. The increase continues until $\approx 33 - 45\%$ of

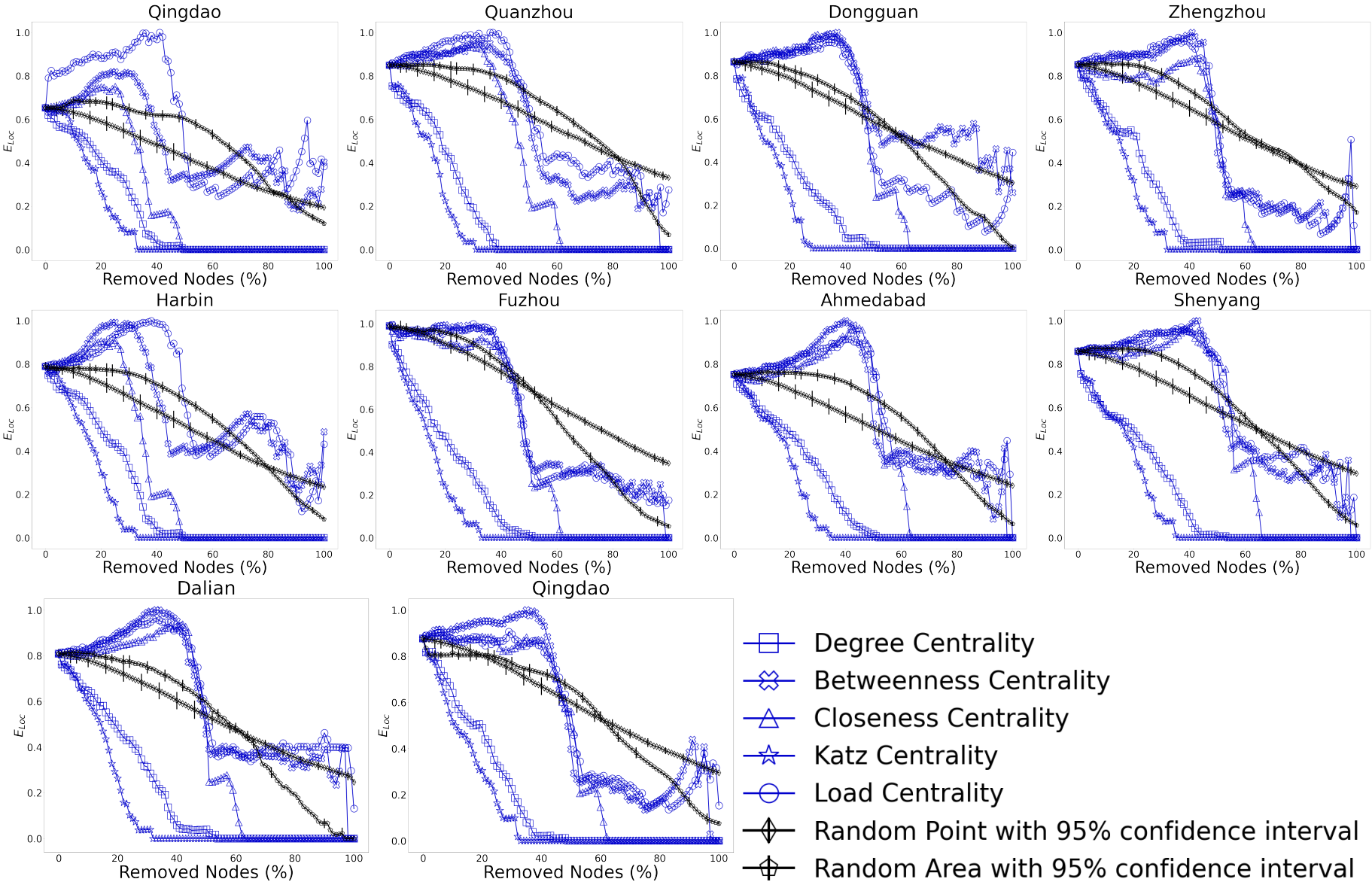


Figure 4.15: The evolution of E_{Loc} curve (normalised) for different disruption strategies in the ten road networks.

the nodes removed from the network (e.g., the E_{Loc} increases for 0.05 in Fuzhou and 0.25 in Surat when removing nodes based on the highest ranked c_L). However, when more than $\approx 45\%$ of the nodes are removed, the network begins to lose its structural integrity and ability to maintain efficient paths between nodes.

We illustrate the possibility of such counter-intuitive phenomenon in Figure 4.16 with a small 12-node network. In this illustration, when node 1 is removed, we see the average E_{Loc} is increased from 0.411 to 0.485. Specifically, we see the E_{Loc} of node 0 is increased from 0.33 to 1.00 due to the removal of node 1 causing the resulting neighborhood of node 0 to consist only of node 2 and 6 which are directly connected.

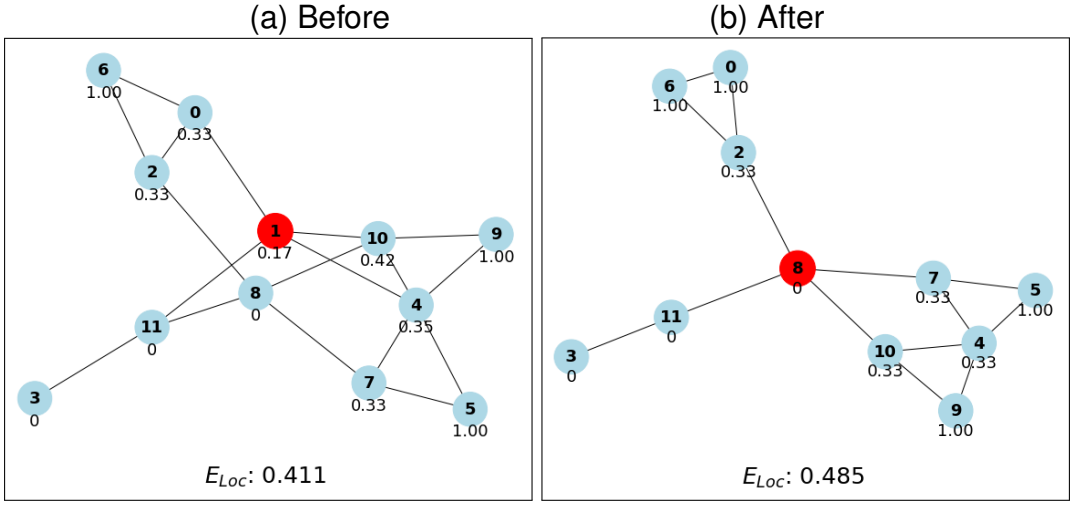


Figure 4.16: Illustration of a node removal increasing E_{Loc} in a sample 12-node network. In this case, removing node 1 increases E_{Loc} from 0.411 to 0.485 (personal collection 2024).

Referring to the definitions of the c_B, c_L, and c_C (Table 4.1), perturbations based on these centralities decrease the number of shortest paths, which in turn, can lead to an increase in network connectivity within local subgraphs. Hence, we observe the initial increase pattern in our results. However, the E_{Loc} eventually falls faster than the random perturbation strategies after the initial increase. In short, we find $(c_K, c_D) \succ (R_P, R_A) \succ (c_B, c_L, c_C)$.

Figure 4.17 shows the percentage of nodes needed to be removed to reduce E_{Loc} by 25%, 50%, and 90%. In general, more nodes are needed to be removed

to achieve the equivalent level of reduction of E_{Loc} than S_{LCC} and E_{Glob}. The c_K node removal strategy appears to be the fastest in disrupting the road networks, followed by c_D.

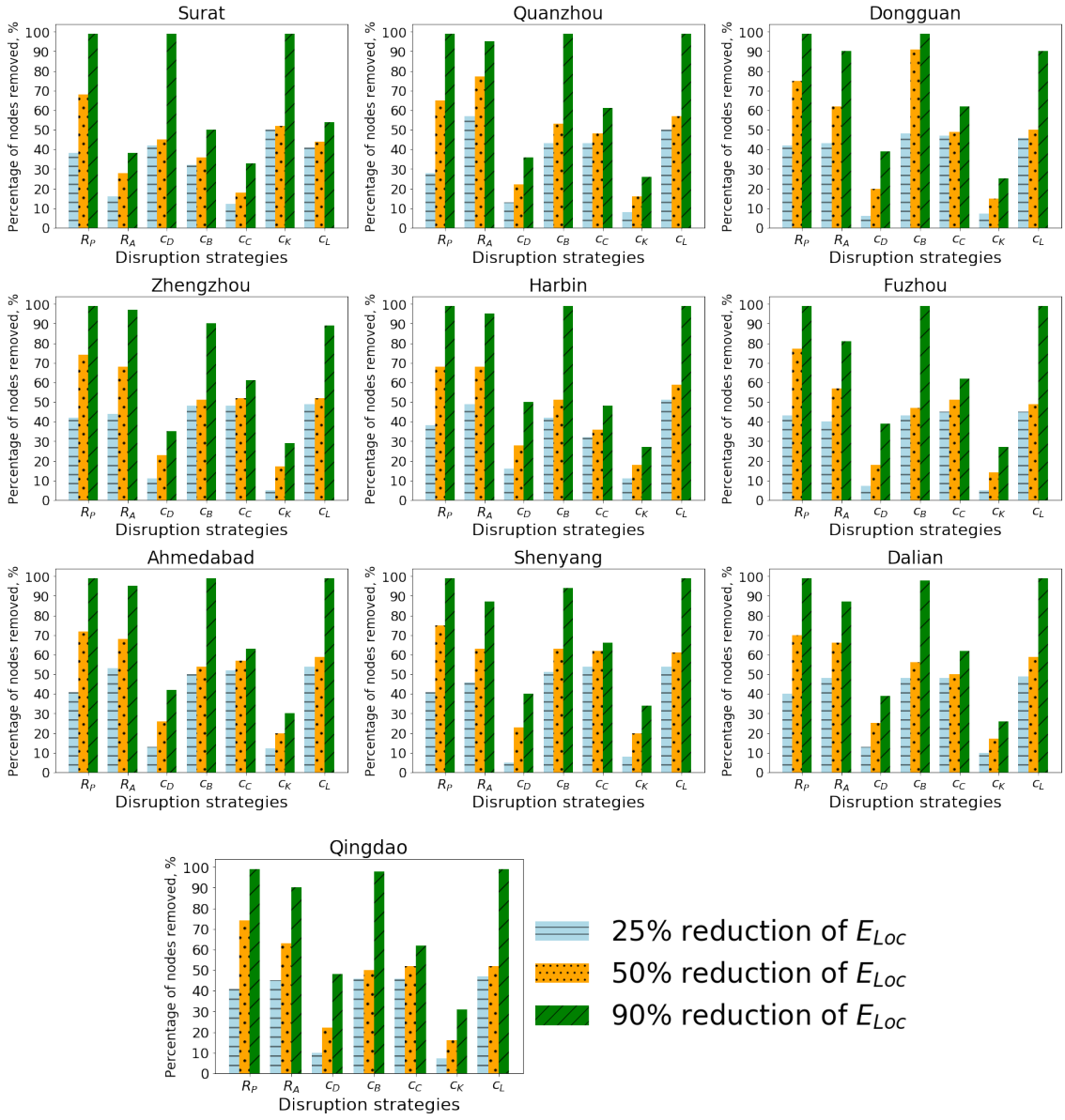


Figure 4.17: Reduction of E_{Loc} based on types of perturbation strategies to achieve 25%, 50%, and 90% decrease.

In order to provide a comprehensive comparison of the networks' response to the considered robustness metrics to the different node removal strategies, we present the scatter plots of the normalised metrics in Figure 4.18, Figure 4.19 and Figure 4.20. In these figures, the red bisector line indicates a perfect correlation between the two robustness metrics (i.e., the network response is the same for

both metrics).

From the Figure 4.18, we see a relatively good correlation between S_{LCC} and E_{Glob} with c_D and c_C mostly appearing above the bisector line indicating that there is a higher decrease of S_{LCC} than E_Glob while the rest recorded the opposite (i.e., faster degradation of E_Glob). We see that R_P and R_A are the closest to the bisector line (Figure 4.18), implying a higher correlation between S_{LCC} with E_{Glob} based random removals. On the other hand, the relationship between S_{LCC} and E_{Loc} showed a weak correlation with sharper S_{LCC} decrease (Figure 4.19). This is most apparent for c_B, c_L, and c_C and mainly due to the initial increase of E_{Loc} discussed above. Similarly, the comparison of E_{Glob} vs E_{Loc} pair indicates a faster decline of E_{Glob} (Figure 4.20).

4.5 Chapter Remarks

In this research, we investigate the robustness of urban road networks in densely populated cities. We use real-world data of these networks and conduct the robustness assessment via an iterative node removal process, monitoring the degradation of the network in terms of the size of the largest connected component, global efficiency, and local efficiency. We considered seven node removal strategies; two of which are stochastic in nature based on random selection and five are deterministic where nodes are ranked based on different centrality measures, namely degree, betweenness, closeness, Katz, and load centralities.

Our results show that the introduction of an increasing amount of perturbations degrades the considered robustness metrics but in different magnitudes. For the size of the largest connected component and global efficiency, random disruption strategies are almost always the least damaging compared to targeted disruption strategies based on centrality measures. We found that random area disruption (where a random neighborhood of the network is gradually disrupted) inflicts higher degradation of S_{LCC} and E_{Glob}. Among the different targeted disruption strategies, we found c_B to be the most disruptive while c_K the least though the difference in effectiveness is less for E_{Glob} than S_{LCC}.

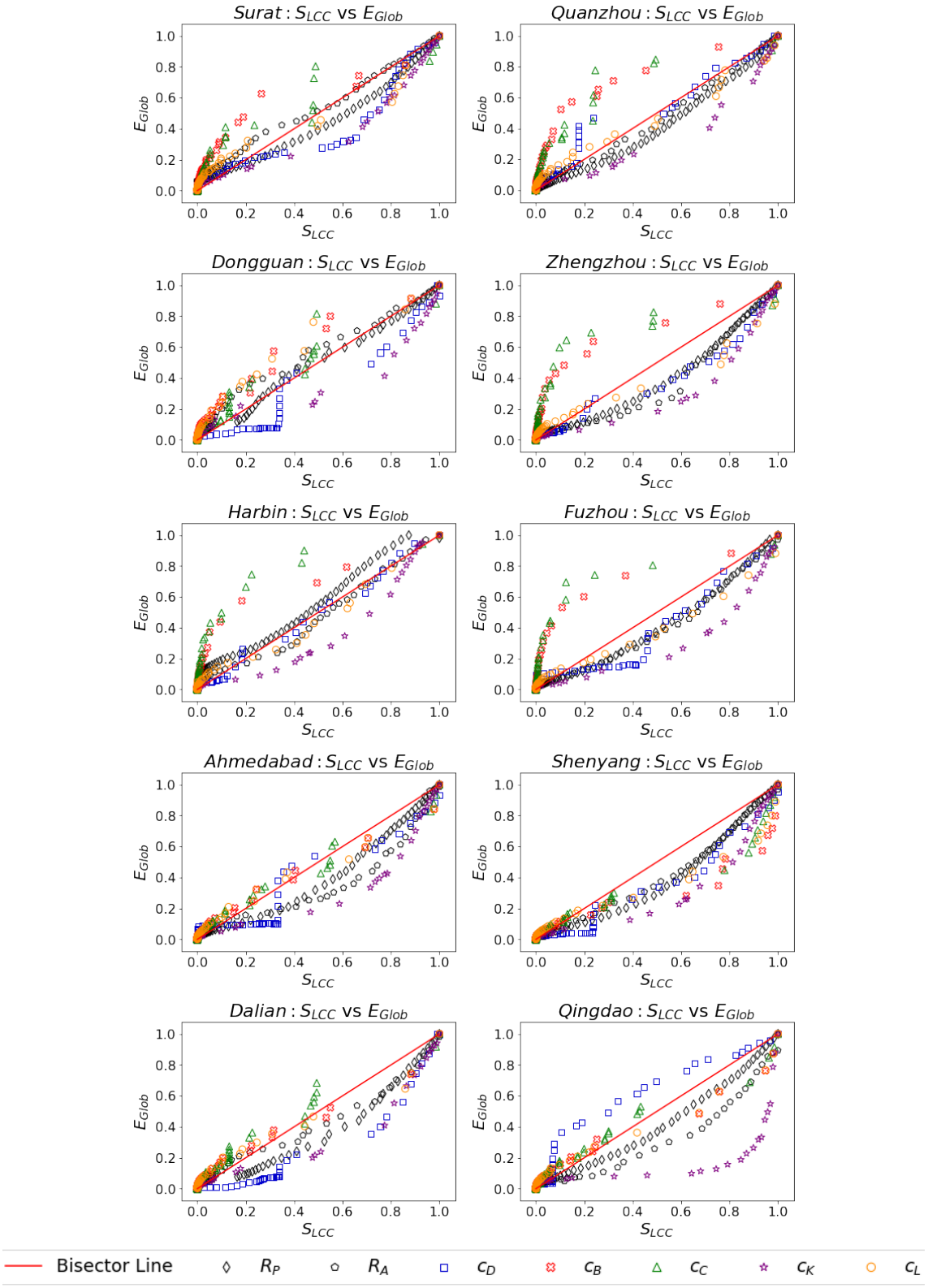


Figure 4.18: Networks functioning comparison. The bisector line indicates the perfect correlation between the two metrics S_{LCC} and E_{Glob}.

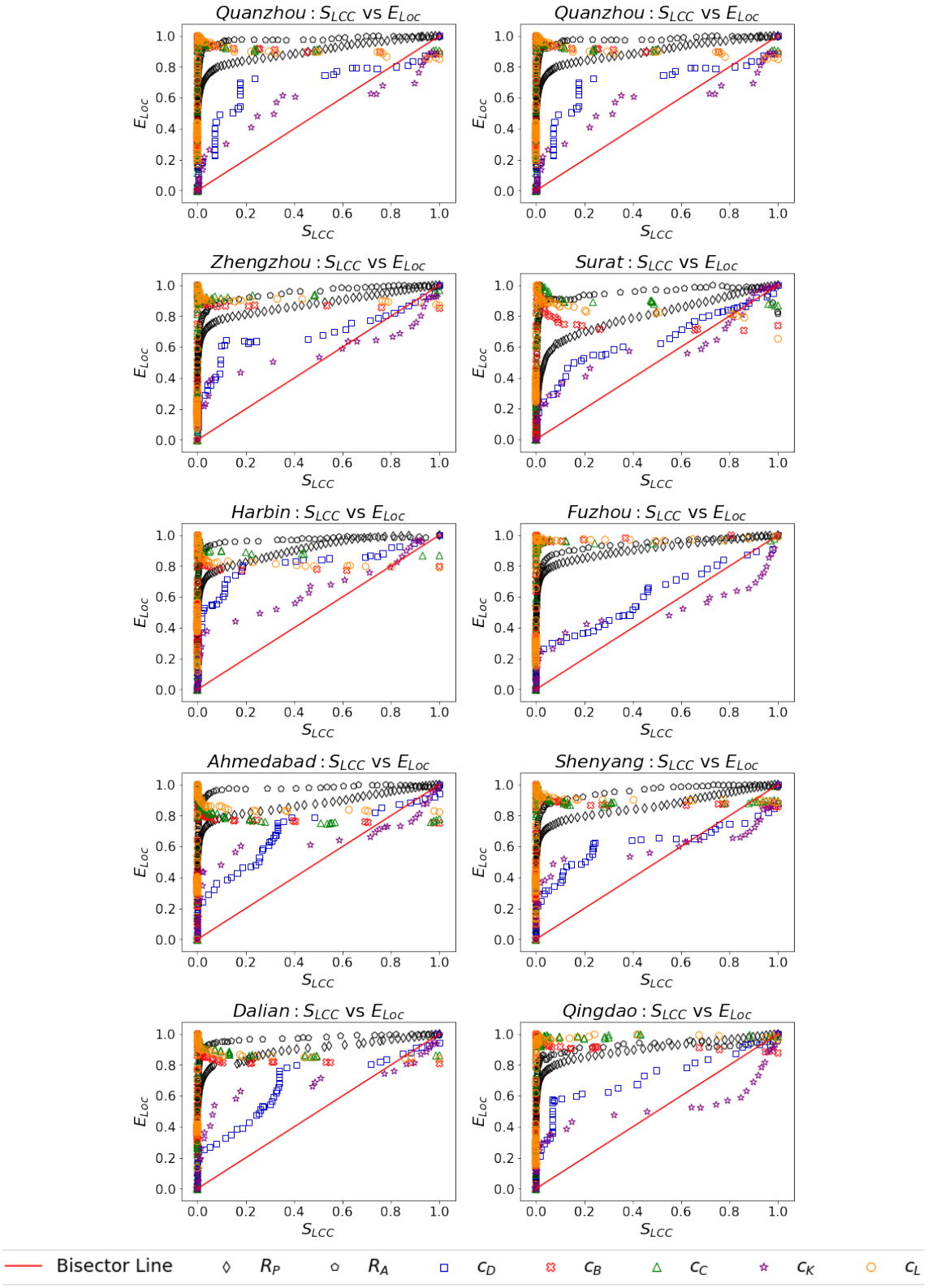


Figure 4.19: Networks functioning comparison. The bisector line indicates the perfect correlation between the two metrics S_{LCC} and E_{Loc}.

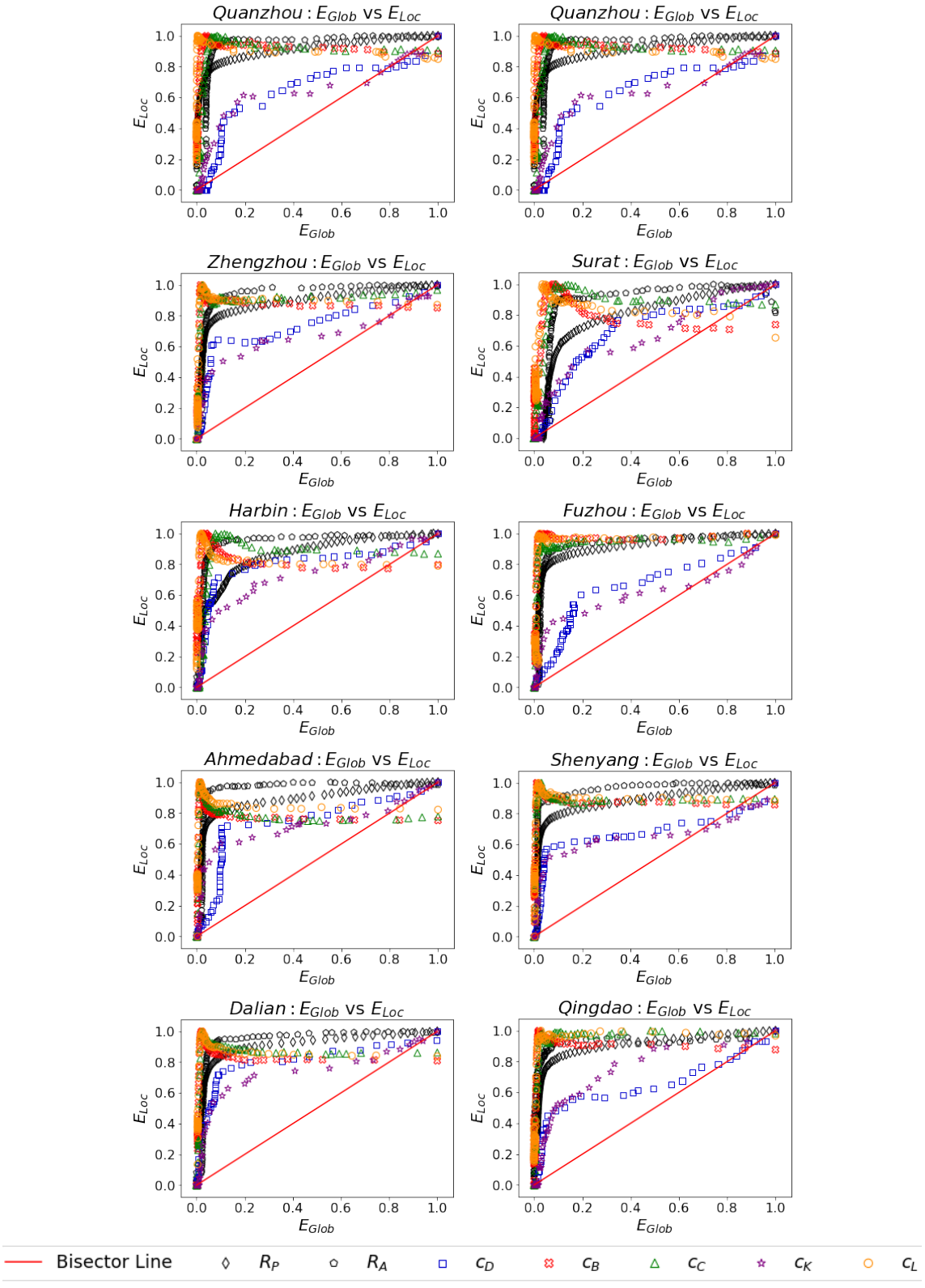


Figure 4.20: Networks functioning comparison. The bisector line indicates the perfect correlation between the two metrics E_{Glob} and E_{Loc}.

While the centrality-based disruption strategies are more effective, they impact the network differently for different robustness metrics. From our analysis, removing nodes with the highest betweenness centrality appears to be the most damaging while c_D and c_K are the least effective according to S_{LCC} and E_{Glob}. Meanwhile, based on our similarity analysis, some disruption strategies correlates with each other. As such, we suggest the combination of c_C with c_K, and c_D with c_C strike a good balance between differences in disruption sequences.

The degradation of local efficiency offers a different picture where we find interestingly an initial increase after perturbations for removal strategies based on c_B, c_C and c_L. From our results, c_K, and c_D inflict greater degradation to local efficiency than other strategic disruption strategies while stochastic removal strategies (i.e., R_P and R_A) lie in between.

Synthesizing the results, for assessing the robustness of road networks, we recommend using c_B as the first choice for studying worst case scenario, R_P and R_A to gain insights into average robustness degradation behavior and finally, using c_D and c_K for finding the impact of the localised impact of perturbations. Considering the entire network as a complex system, removal strategies based on centrality that takes into account paths are more disruptive.

Chapter 5

Modeling traffic congestion spreading using a topology - based SIR epidemic model

In the realm of urban road networks and traffic congestion, dynamic processes involve the changes in behaviors and patterns within road systems as time progresses. The spread of congestion, akin to an epidemic, is a significant component of this phenomenon. Researchers employ the Susceptible - Infected - Recovered (SIR) model, commonly used in epidemiology, to analyse and forecast the spread of congestion in urban settings.

Through incorporating the road network structure into the model using a N - intertwined framework, the research aims to provide a deeper understanding of congestion propagation dynamics. Real-world traffic data from California and Los Angeles are utilised to validate the proposed model against traditional SIR models, yielding promising results that align closely with actual congestion patterns. This work not only enhances our comprehension of congestion diffusion but also offers valuable insights for the development of effective congestion mitigation strategies.

5.1 Modeling Traffic Congestion Spreading with SIR Topology-Awareness

Urban traffic commonly demonstrates significant spatial correlation, where nodes adjacent to congested areas are more prone to congestion. Moreover, there exists a notable temporal correlation in urban congestion driven by varying travel demand patterns over time. While congestion dynamics at the link-level are well understood through queuing and kinematic wave theories, our comprehension of congestion propagation mechanisms at the network level remains incomplete.

Queue spillbacks in networks appear sensitive to link capacities, remaining stable under both under- and over-saturated conditions. During recovery, congestion tends to fragment, resulting in increased spatial heterogeneity and reduced network efficiency. The propagation and dissipation of congestions can be assessed by the number or length of congested links in the network, with propagation typically occurring at a faster rate than dissipation. Studies indicate that the road supply-to-travel demand ratio can explain the percentage of time lost in congestion as a comprehensive measure

As previously mentioned, there are three fundamental contagious models of infectious diseases: SI, SIS, and SIR. Contagion processes, originally inspired by how diseases spread, depict the interaction between those susceptible to infection and those already infected across a network of connections. However, this concept extends beyond just disease transmission. It is also utilised to explain the spread of ideas on social media, the proliferation of computer viruses online, and the formation of traffic congestion in urban areas. One prevalent method for analysing contagion is compartmental modeling, where each element in the network is categorised into specific states or compartments.

Various contagion models, like susceptible-infected (SI) and susceptible-infected-susceptible (SIS), can also be applied to represent traffic spread in a network from a two-class perspective. The SI model may illustrate congestion propagation as long as travel demand persists, leading eventually to network-wide gridlock

without recovery. Conversely, the SIS model can describe congestion dynamics where the traffic jam grows but does not encompass the entire network, maintaining a partial blockage over time. While both SI and SIS models theoretically capture congestion spread, their accuracy remains uncertain.

In situations involving fluctuating congestion within a network, the SIR model, as a three-class model, offers a more practical representation, supported by empirical and simulation-based evidence. From the literature, the SIR model has been adopted as the suitable model for replicating congestion propagation and provides a realistic representation (Wu et al. 2004; Brockmann and Helbing 2013; Saberi et al. 2020). As such, we also adopt the SIR epidemic model in this work.

5.1.1 Preliminary

The SIR model (Kermack and McKendrick 1927) is widely used in epidemiology. By compartmentalisation, a disease is broken down into three distinct stages, namely Susceptible (state S) to represent healthy individuals, Infected (state I) to represent infected individuals who are also Infectious, and Removed (state R) to represent an individual who has recovered from an infection. As indicated in Chapter 2, this model has gained significant attention in the field of traffic congestion spreading due to its simplicity and performance in capturing congestion dynamics. In the context of this research, the SIR model is applied to represent the following states:

- Susceptible (S) – nodes that are not congested (free flow traffic) but susceptible to be congested,
- Infected (I) – nodes that are congested, and the congestion may overspill to the neighbouring nodes,
- Recovered (R) – nodes that suffered from congestion, but the congestion has since dissipated.

Using the law of mass action, the SIR model could be modeled with the following system of ordinary differential equations (ODEs) (Kermack and McKendrick 1927;

Anderson and May 1991):

$$\frac{ds(t)}{dt} = -\beta s(t)i(t) \tag{5.1}$$

$$\frac{di(t)}{dt} = \beta s(t)i(t) - \gamma i(t) \tag{5.2}$$

$$\frac{dr(t)}{dt} = \gamma i(t) \tag{5.3}$$

where $s(t)$, $i(t)$ and $r(t)$ denote the number of nodes in susceptible, infected, and recovered state at time t respectively while β and γ are the propagation and dissipation rates. Hereafter, in this research, we refer to the above model as the *classical SIR* model.

Using the homogeneous mixing hypothesis, the classical SIR model can further consider including the average contacts of each node in a network into the system of equations as follows (Barabási 2016):

$$\frac{ds(t)}{dt} = - < k > \beta s(t)i(t) \tag{5.4}$$

$$\frac{di(t)}{dt} = < k > \beta s(t)i(t) - \gamma i(t) \tag{5.5}$$

$$\frac{dr(t)}{dt} = \gamma i(t) \tag{5.6}$$

where $< k >$ denotes the average degree of the network. Hereafter, we refer to this model as *average-degree-based SIR* model. Both the above models neglect the heterogeneity of nodes and thus, have not considered the network structure.

5.1.2 Topology-based SIR Model

To account for the heterogeneity of the individual road segments in a road network, we adopt the modeling framework proposed in Mieghem et al. (2008), named the N-intertwined epidemic model. The framework employs a continuous-time Markov chain analysis to model spreading behavior. It has been studied and extended in various directions (e.g., Mieghem 2011; Youssef and Scoglio 2011; Sahneh et al. 2013; Chai and Pavlou 2017; Chai 2017). The framework is general and applicable to different problem domains. It allows us to study how the underlying network structure affects congestion spreading patterns.

In the context of our problem, consider a road network represented by an undirected graph, $G(V, E)$, where V and E are the set of N nodes and L links. In our case, the nodes represent the sensors (e.g., inductive loops, roadside traffic cameras, etc.) deployed at the roads to monitor traffic while links represent road segments connecting neighbouring sensors. Graph, G, can be represented by A, the $N \times N$ adjacency matrix, with $a_{n,m} = 1$ if there exists a link between nodes n and m, and 0 otherwise.

We say a node gets infected when a node previously having free-flowing traffic becomes congested (i.e., a transition from state S to state I). Consistent with the previous models introduced in Section 5.1.1, this infection process takes place at an average rate of β per unit of time. Similarly, a node is said to have recovered from congestion (i.e., a transition from state I to state R) when the traffic returns to a normal free-flowing state (see Figure 5.1). Again, we use γ as the average recovery rate per unit of time.

Next, we formally define the notion of "congestion". Given the average vehicle speed recorded at node n between time t and $t - \Delta$, $v_n(t)$ where Δ is the time interval dependent on the dataset (e.g., 5-min) and the speed limit at node n, v_n^{max}, we define the following:

$$\lambda_n(t) = \frac{v_n(t)}{v_n^{max}} \tag{5.7}$$

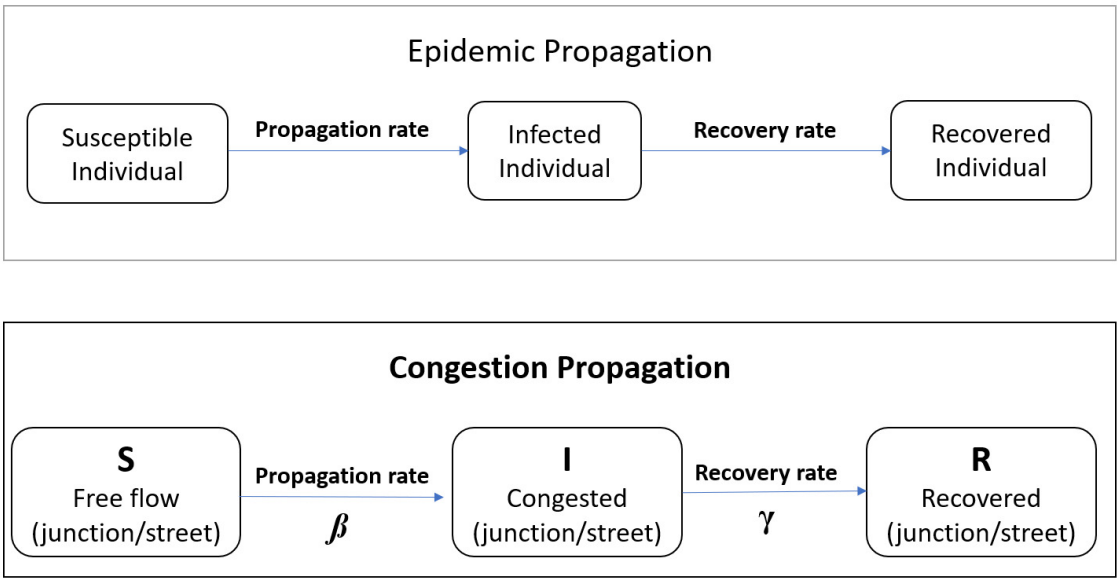


Figure 5.1: Physical mapping of congestion spreading on the SIR model (Kozhabek et al. 2024).

as the measure of traffic flow at the node n. A low value of $\lambda_n(t)$ means the vehicle is moving slowly or non-moving if $\lambda_n(t) = 0$. Conversely, the traffic is free-flowing when $\lambda_n(t)$ is high. Node n is considered as congested (i.e., "infected" in the epidemiology terminology) if $\lambda_i(t) < \rho$ where ρ represents different congestion levels (Li et al. 2015). Otherwise, a node is considered not congested (i.e., $\lambda_i(t) \geq \rho$). The number of congested nodes grows as ρ increases. ρ is a tuneable threshold value (cf. Section 5.3 on the impact of different values of ρ).

Let $s_n(t)$, $i_n(t)$, and $r_n(t)$ denote the probability of node n being in the susceptible, infected, and recovered state at time, t, respectively. Since each node can be in one and only one of the three possible states at any one time, then we have:

$$s_n(t) + i_n(t) + r_n(t) = 1 \tag{5.8}$$

and

$$\frac{ds_n(t)}{dt} + \frac{di_n(t)}{dt} + \frac{dr_n(t)}{dt} = 0. \tag{5.9}$$

The complexity of the solution by applying the Markov theory directly to the entire network is exponential (i.e., $O(3^N)$) since all possible combinations of states for

each and every node have to be considered. This results in the infinitesimal generator of the system, $Q(t)$ having the dimension of $3^N \times 3^N$. To address this issue, we advocate the use of the N-intertwined epidemic framework (Mieghem 2011; Youssef and Scoglio 2011; Sahneh et al. 2013; Chai and Pavlou 2017; Chai 2017) for the SIR model which approaches the problem by considering each node individually. Now, applying the Markov theory, we will get N infinitesimal generators, $Q_n(t)$ of the three-state continuous Markov chain; one for each node as follows:

$$Q_n(t) = \begin{bmatrix} -q_{1,2;n} & q_{1,2;n} & 0 \\ 0 & -q_{2,3;n} & q_{2,3;n} \\ 0 & 0 & 0 \end{bmatrix} \tag{5.10}$$

where $q_{1,2;n}$ is dependent on the states of other nodes within the network. One way to account for this dependency is to condition $q_{1,2;n}$ with all possible combinations of states for all nodes. However, this reverts back to the exact Markov chain solution of exponential complexity. Hence, we apply a mean field approximation to the random variable $q_{1,2;n}$ and compute its expected rate. By this, we remove the random nature and it allows us to reduce the complexity of the solution to a polynomial ($O(N)$).

The effective infinitesimal generator, $\overline{Q_n(t)}$, is then as follows:

$$\overline{Q_n(t)} = \begin{bmatrix} -E[q_{1,2;n}] & E[q_{1,2;n}] & 0 \\ 0 & -\gamma & \gamma \\ 0 & 0 & 0 \end{bmatrix} \tag{5.11}$$

where $E[q_{1,2;n}] = \sum_{m=1}^{N} a_{m,n} i_n(t)$.

We can now obtain the following system of non-linear governing differential equations for our case,

$$\frac{ds_n(t)}{dt} = -s_n(t)\beta \sum_{m=1}^{N} a_{m,n} i_n(t) \tag{5.12}$$

$$\frac{di_n(t)}{dt} = s_n(t)\beta \sum_{m=1}^{N} a_{m,n} i_n(t) - \gamma i_n(t) \tag{5.13}$$

$$\frac{dr_n(t)}{dt} = \gamma i_n(t) \tag{5.14}$$

Note the explicit consideration on the road network topology via the inclusion of the adjacency matrix elements in the system. This allows us to validate the model against real road networks and congestion data (cf. Section 5.3). Another advantage of adopting this framework is that we can now consider the state transitions of each node individually, as opposed to the classical methods where the mean aggregate behaviour of all nodes is considered.

We can now solve the above system of differential equations to obtain the instantaneous evolution of the nodes for the three distinct states to discover the congestion spreading patterns of road networks. By additionally using Eqs. 8 and 9, we can reduce the problem from solving $3 \times N$ simultaneous equations to $2 \times N$ equations.

Following the above, we can then obtain the epidemic prevalence (i.e., the fraction of nodes in congested state (infected)) at time t as follows:

$$P(t) = \frac{1}{N} \sum_{n=1}^{N} i_n(t). \tag{5.15}$$

At steady-state ($\frac{di_n(t)}{dt}|_{t\to\infty} = 0$), $i_{n\infty} \equiv \lim_{t\to\infty} i_n(t) = 0$ and consequently, $\sum_n^N s_n + \sum_n^N r_n = N$.

5.2 Data for Traffic Congestion Modeling

The road networks used earlier, as outlined in Section 3.4, do not contain traffic data. Our focus in this chapter is on modeling the diffusion of traffic congestion

by utilising alternate urban road network datasets with traffic speed information. In particular, we make use of two real-world datasets named PEMS-BAY and METR-LA sourced from Li et al. (2017a). These datasets offer crucial insights enabling us to study and simulate the dynamics of traffic congestion in urban settings. The PEMS-BAY dataset is collected from the California Transportation Agencies (CalTrans) Performance Measurement System (PeMS) (Chen 2002). It recorded traffic speed data from 325 sensor stations in the Bay Area.

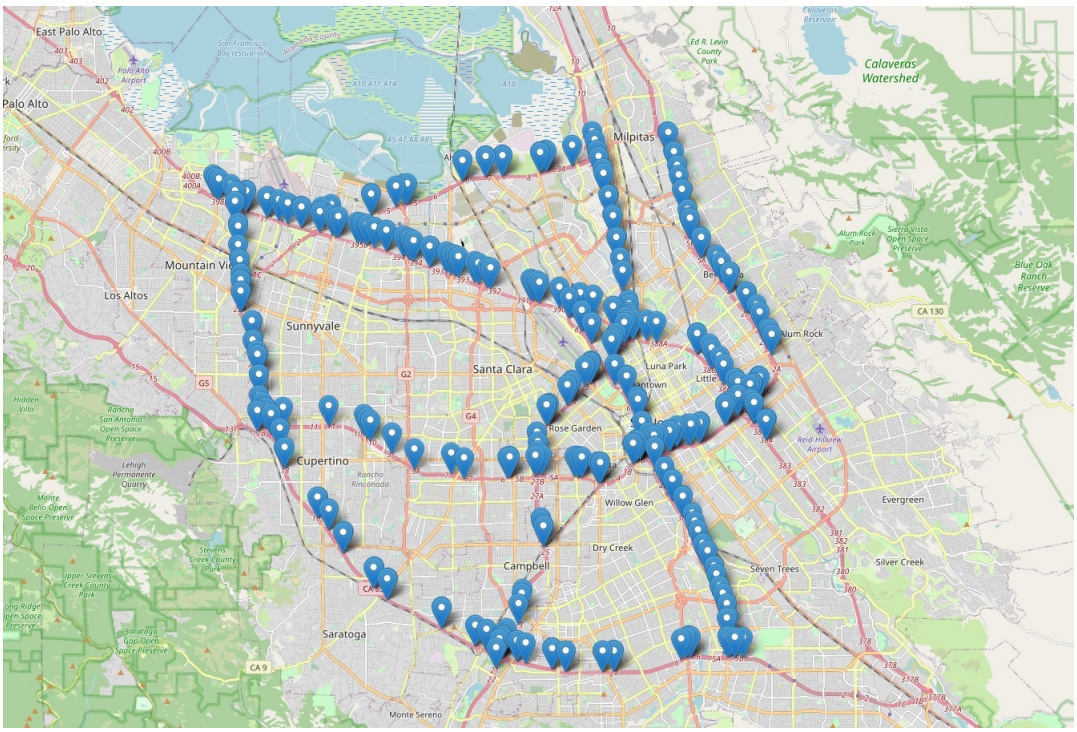


Figure 5.2: Sensor distribution of the PEMS-BAY dataset

The locations of the sensor stations are shown in Figure 5.2. The traffic measurements cover a period of six months between the 1st of January and the 30th of June in 2017. The time interval for the data is 5 minutes and the total number of observed traffic data points is 16,937,700 ($= 52,116 \times 325$).

The second real-world dataset, METR-LA, shown in Figure 5.3, was collected from loop detectors in the highways of Los Angeles County (Jagadish et al. 2014). It includes 207 sensors and covers traffic data for a period of four months from the 1st of March to the 30th of June in 2012. The time interval between data points is also 5 minutes, and the total number of observed traffic data points is 7,094,304 ($= 34,272 \times 207$) with 8.11% missed data points due to sensor incidents.

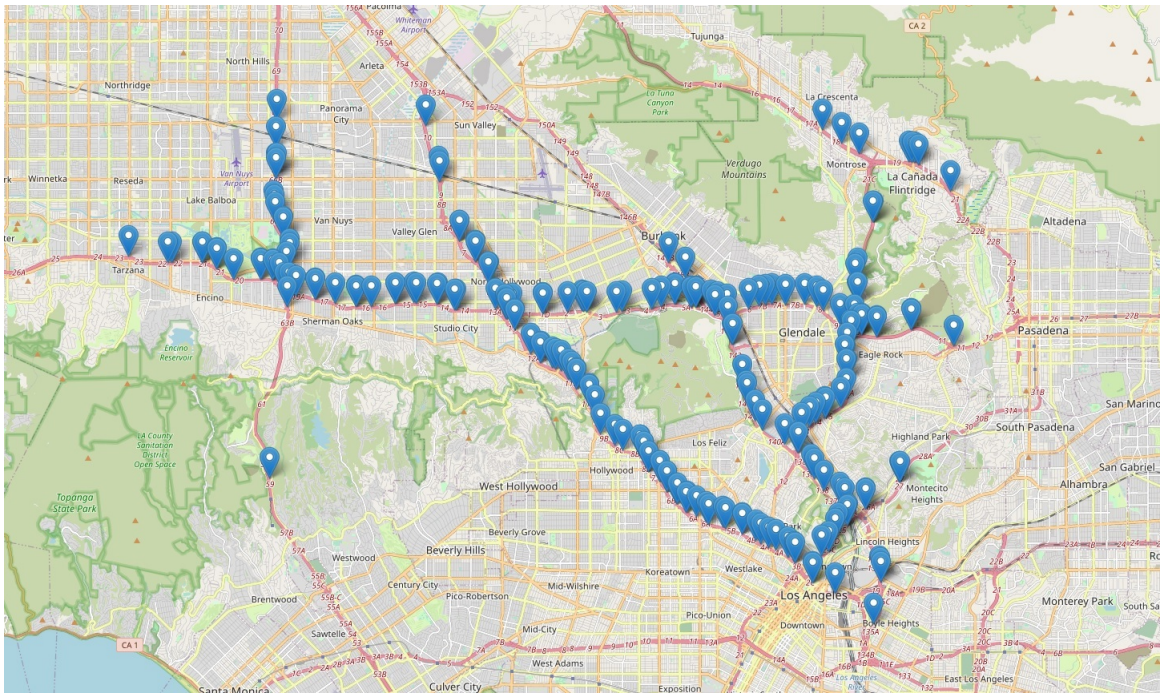


Figure 5.3: Sensor distribution of the METR-LA dataset

Table 5.1 provides a summary of the basic statistics of both datasets including maximum (Max), minimum (Min), mean value (Mean), standard deviation (Std), and variance (Var) of traffic speed data as well as the size of the dataset (in MByte). From the table, it can be noted that METR-LA has higher traffic fluctuations with larger standard deviation and variance than PEMS-BAY.

Table 5.1: Traffic speed characteristics of datasets

Datasets	Max	Min	Mean	Std	Var	Size
PEMS-BAY	85.1	0.0	62.62	8.56	85.41	136 MB
METR-LA	70.0	0.0	53.72	19.2	374.9	57 MB

Table 5.2: Characteristics of adjacency matrices

Datasets	Nodes	Edges	Average node degree
PEMS-BAY	325	398	2.098
METR-LA	207	264	2.058

Since we advocate the consideration of the actual road topology, we also constructed the adjacency matrix for both datasets based on road connectivity. The

basic characteristics of the road network topology derived from their respective adjacency matrices are presented in Table 5.2.

5.2.1 Topological analysis: PEMS-BAY and METR-LA

In this subsection, we perform a multi-scale topological analysis of road networks, as introduced in Chapter 3. It should be underlined that the PEMS-BAY and METR-LA road networks were abstracted differently from the dataset used in previous chapters, resulting in significant differences in size and detail. PEMS-BAY consists of 325 nodes and 398 edges, while METR-LA includes 207 nodes and 264 edges, both focusing on major roads and intersections with a coarse-grained abstraction. In contrast, the dataset for Surat, used in Chapters 3 and 4, is much larger and more detailed, with 2593 nodes and 7340 edges, capturing not only major roads but also smaller local roads and intersections. The difference in abstraction levels highlights the varying scales and scopes of these datasets, which must be considered when comparing results across them.

Nevertheless, as shown in Table 5.3 and Table 5.4, we repeated the topological analysis conducted in Chapter 3 for consistency and comparison. The values between the macro-scale analysis in Table 5.3 and the analysis in Table 3.2 differ, reflecting the varying levels of abstraction in the studied road networks. Specifically, the values for PEMS-BAY and METR-LA are lower across all nine macro metrics.

Table 5.3: Macro-scale metrics results for PEMS-BAY and METR-LA

Road networks	r^T	m	$\bar{d}$	κ	CC_G	τ_G	D	H_G	eff_G
PEMS-BAY	0.00033	0.115	2.1	2.2	0.016	0.031	52	21.47	0.063
METR-LA	0.01969	0.142	2.3	2.6	0.048	0.055	37	15.11	0.106

Table 5.4: Meso-scale metrics results PEMS-BAY and METR-LA

Road networks	Modularity, Q	Number of communities	Mean community size	*Path-Core*,%
PEMS-BAY	0.865	18	17	21.25
METR-LA	0.803	12	17.08	17.85

5.2.2 Robustness assessment: PEMS-BAY and METR-LA

In this subsection, we evaluate the robustness of the PEMS-BAY and METR-LA road networks under five targeted perturbation strategies and two random perturbation scenarios, following the methodology outlined in Chapter 4. The assessment begins with an analysis of the largest connected component (LCC), as shown in Figure 5.4. All five targeted strategies perform similarly well, making it difficult to identify the most effective one from the figure. Random perturbations are less effective, which is consistent with the results shown in Figure 4.4. A robustness assessment was also conducted on E_{Glob} (in Figure 5.5), where, similarly, random disruptions proved less effective. Among the targeted strategies, c_B-based disruptions were the most effective, followed by c_L. Finally, based on the E_{Loc} curve in Figure 5.6, there is a slight improvement initially, followed by a decrease in the case of targeted perturbations. For random removal, fluctuations are observed. First, efficiency improves, then decreases, but ultimately improves again.

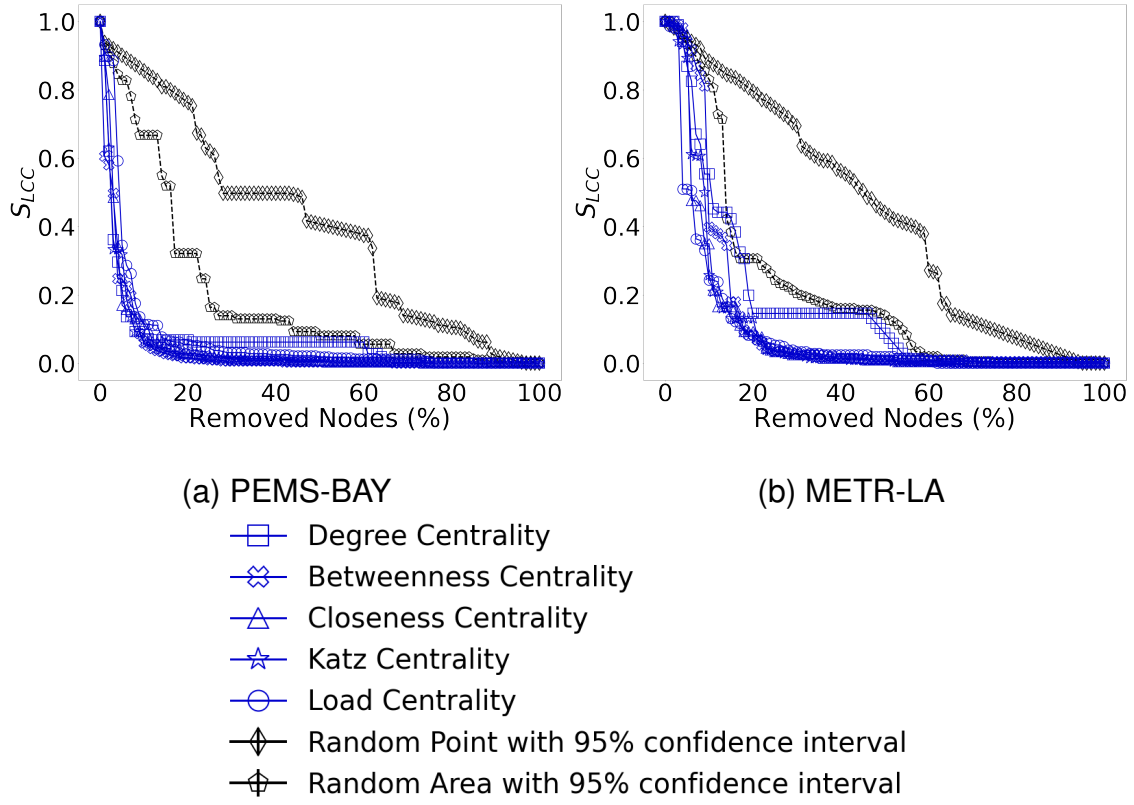


Figure 5.4: The evolution of S_{LCC} curve (normalised) for different disruption strategies in the PEMS-BAY and METR-LA road networks.

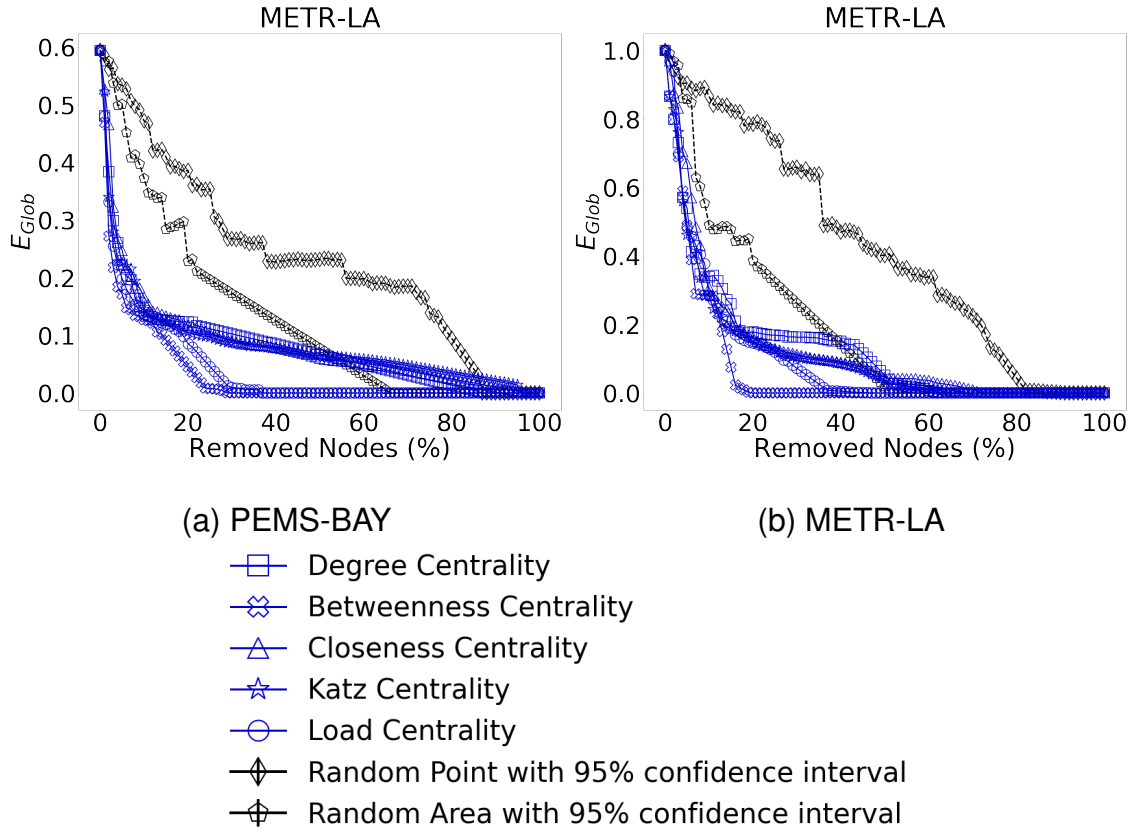


Figure 5.5: The evolution of E_{Glob} curve (normalised) for different disruption strategies in the PEMS-BAY and METR-LA road networks.

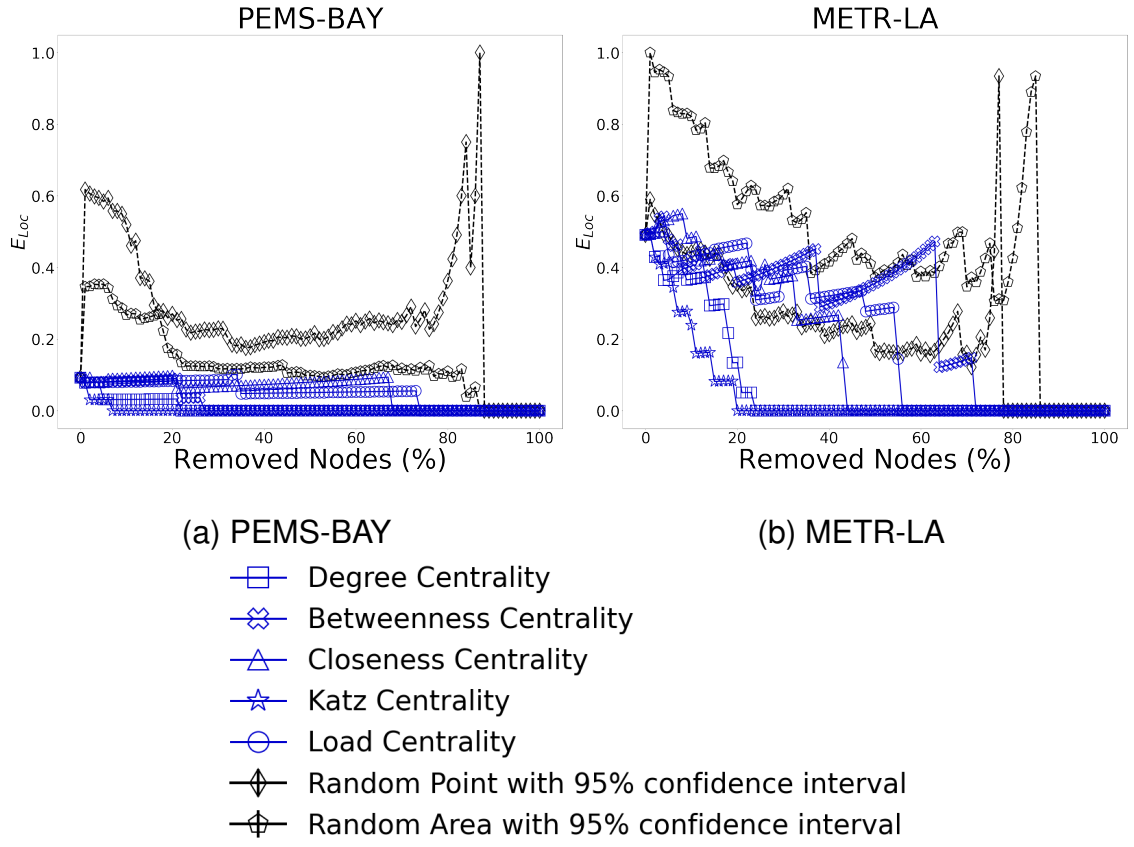


Figure 5.6: The evolution of E_{Loc} curve (normalised) for different disruption strategies in the PEMS-BAY and METR-LA road networks.

5.3 Evaluation

5.3.1 Parameter estimations

To determine the propagation (β) and dissipation (γ) parameters, we follow Marinov et al. (2014) and use an ordinary least squares (OLS) approach along with a pattern search algorithm. The goal of the estimation process is to minimise the root-mean-squared error (RMSE) between the observed and modeled $i(t)$ values over the specified study period.

Table 5.5: Parameters β and γ for different values of ρ for PEMS-BAY.

ρ	β	γ
0.4	1.70	0.45
0.5	2.25	0.70
0.6	2.40	0.80
0.7	2.55	0.95

Parameters for both datasets were outlined in the Table 5.5 for PEMS-BAY and Table 5.6 for METR-LA.

Table 5.6: Parameters β and γ for different values of ρ for METR-LA.

ρ	β	γ
0.4	1.15	0.30
0.5	1.31	0.35
0.6	1.75	0.49
0.7	1.98	0.67

5.3.2 Temporal evolution of traffic congestion

We compare the predictive capacity of the three models, aiming to assess how these models perform over time in tracking the congestion state across the entire network, i.e.,

(1) classical SIR,

(2) average-degree-based SIR,

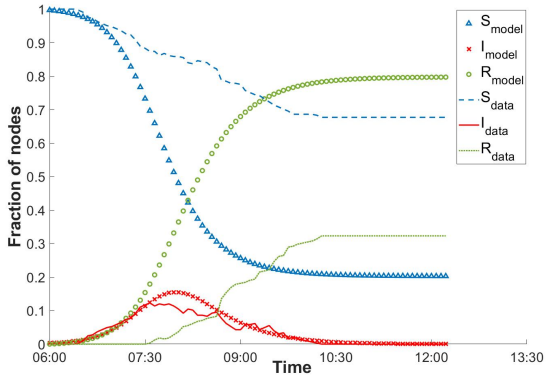
(3) topology-based SIR models, over time.

For this, we first show how the models track the network-wide congestion state focusing on the early peak hour period (i.e., between 06:00 am to 12:00 pm) of a workday.

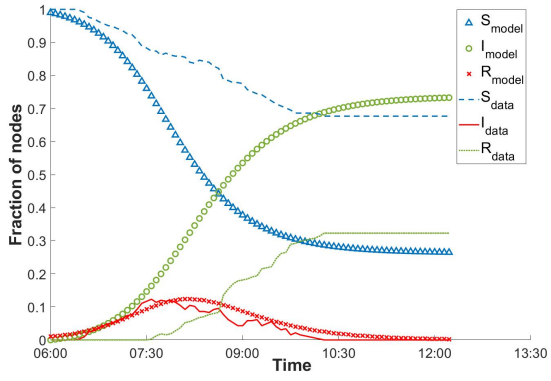
We show in Figure 5.7 for PEMS-BAY and Figure 5.8 for METR-LA the fraction of nodes in different states over the specified time period over three different congestion thresholds (i.e., $\rho = 0.4, 0.5, 0.6$). The S_{data}, I_{data}, and R_{data} curves are the actual traffic recorded in the dataset. Specifically, we present the average traffic speed data taken on all Wednesdays across the measurement period as representative observations. The curves are qualitatively similar for other days of the week.

On the other hand, the S_{model}, I_{model}, and R_{model} curves are computed by the respective models described in Section 5.1. For both figures, the first column includes the performance of the classical SIR model, the second column incorporates the average-degree-based SIR model's outcomes and the third column presents the results of our topology-based SIR model.

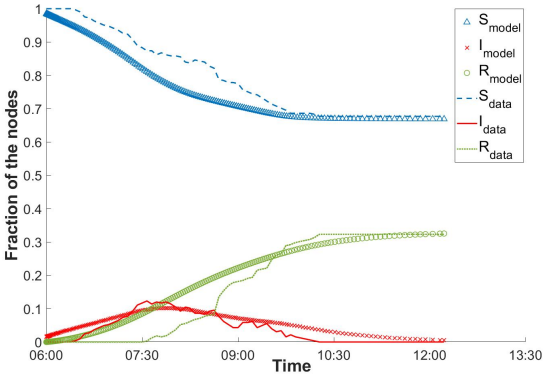
From the figures, we see that, in comparison, the proposed topology-based SIR model captures most accurately all three traffic states over the entire period and can generally track the evolution of the congestion closely for the different ρ considered as well as across both road networks. Taking the PEMS-BAY network at $\rho = 0.4$ as an example, the real traffic data indicates that there are approximately 70% of nodes in the network that never suffered traffic congestion (conversely, the remaining 30% have recovered from congestion at the end of the peak period). The classical SIR model, however, predicted only 20% of nodes avoided congestion during this period (i.e., a 50% discrepancy from the ground truth). The average-degree-based SIR model performed slightly better, computed about 30% of nodes not congested. The topology-based SIR model advocated here managed to correctly predict the outcome at 70% of nodes. Further, despite having



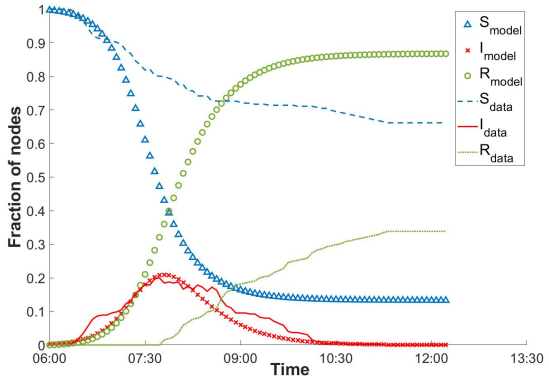
Classical SIR model, $\rho = 0.4$



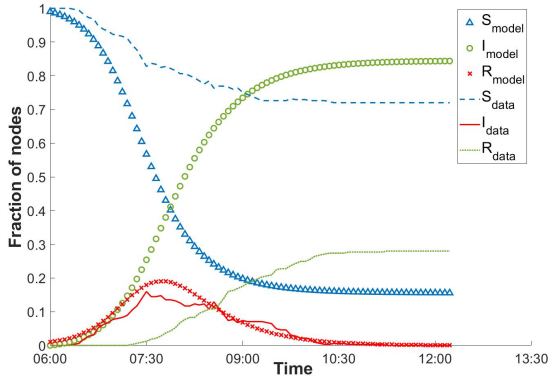
Avr-degree-based SIR model, $\rho = 0.4$



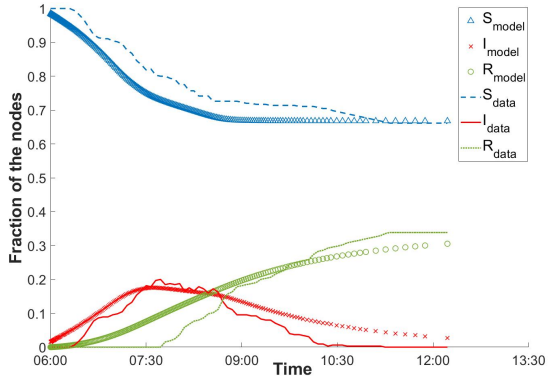
Topology-based SIR model, $\rho = 0.4$



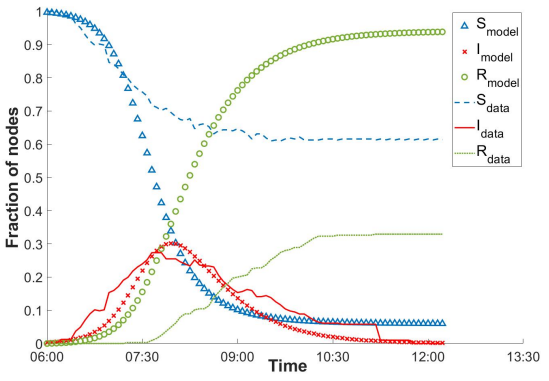
Classical SIR model, $\rho = 0.5$



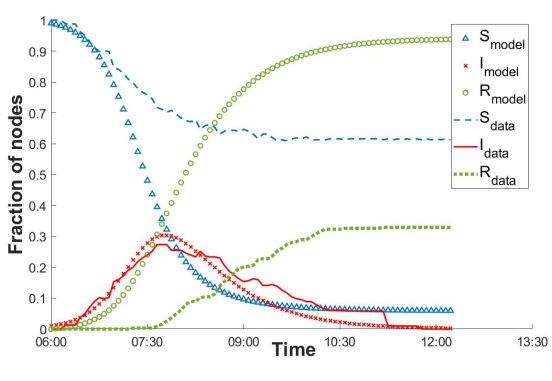
Avr-degree-based SIR model, $\rho = 0.5$



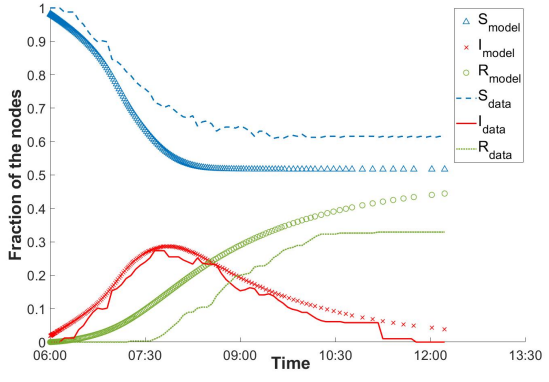
Topology-based SIR model, $\rho = 0.5$



Classical SIR model, $\rho = 0.6$

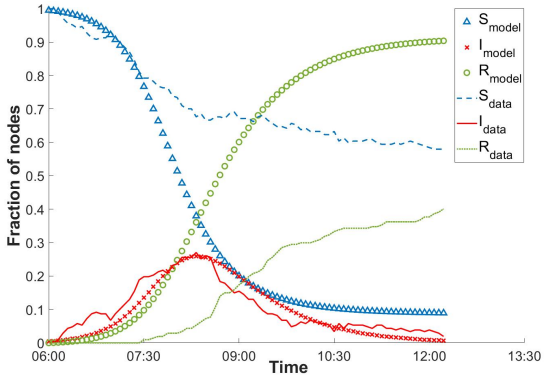


Avr-degree-based SIR model, $\rho = 0.6$

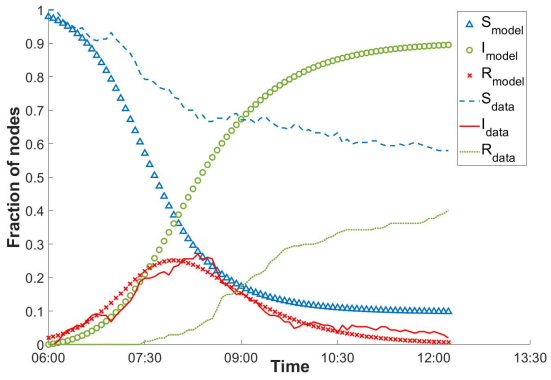


Topology-based SIR model, $\rho = 0.6$

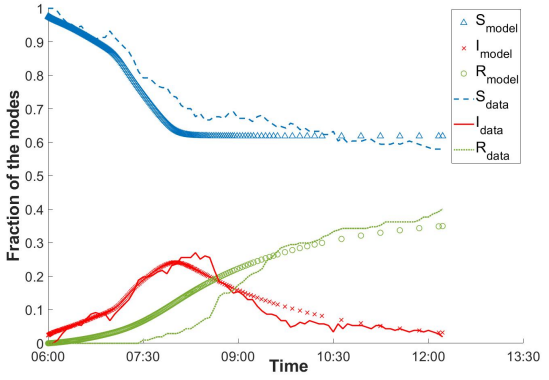
Figure 5.7: Models vs Data – Time evolution of node states for $\rho = 0.4, 0.5, 0.6$ on the PEMS-BAY dataset (Kozhabek et al. 2024).



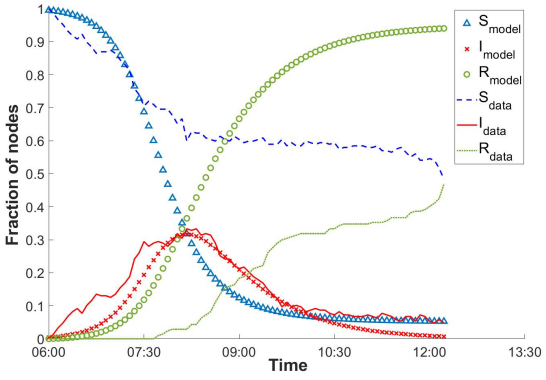
Classical SIR model, $\rho = 0.4$



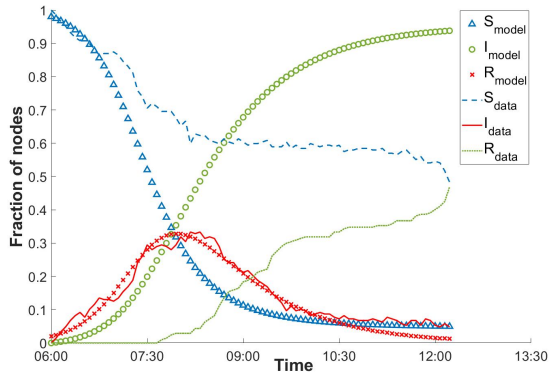
Avr-degree-based SIR model, $\rho = 0.4$



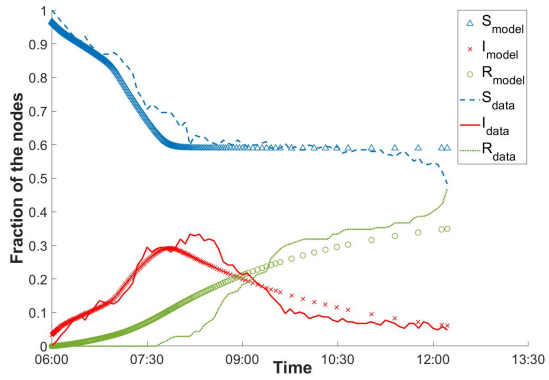
Topology-based SIR model, $\rho = 0.4$



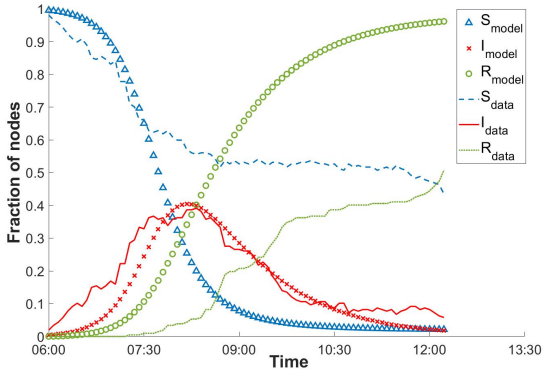
Classical SIR model, $\rho = 0.5$



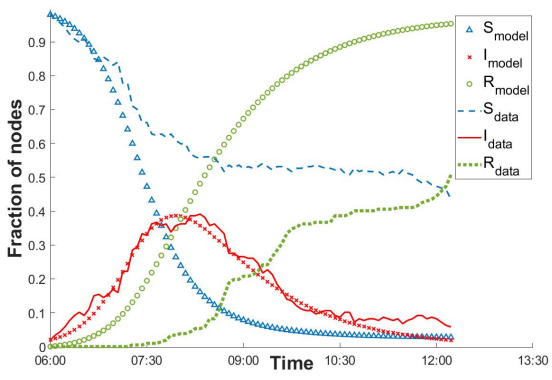
Avr-degree-based SIR model, $\rho = 0.5$



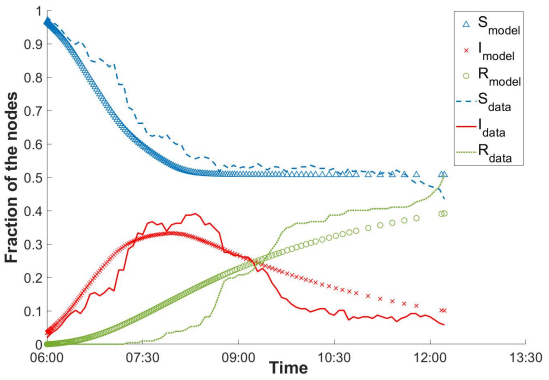
Topology-based SIR model, $\rho = 0.5$



Classical SIR model, $\rho = 0.6$



Avr-degree-based SIR model, $\rho = 0.6$



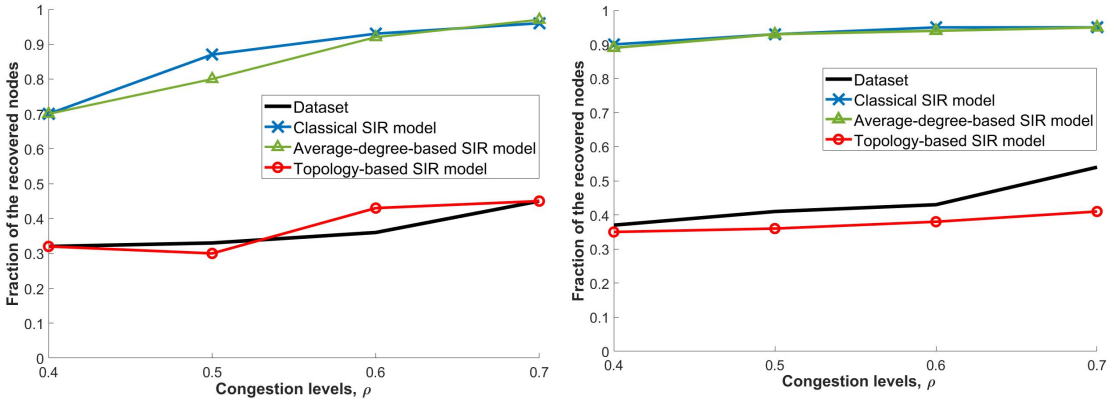
Topology-based SIR model, $\rho = 0.6$

Figure 5.8: Model vs Data – Time evolution of node states for $\rho = 0.4, 0.5, 0.6$ on the METR-LA dataset (Kozhabek et al. 2024).

higher traffic volatility in the METR-LA dataset, our topology-based SIR model can still track the congestion state closely.

By definition, when ρ is small, fewer nodes are considered to be congested (i.e., vehicles must be traveling at a lower speed or being stationary at the junction to be classified as congested). We see this across both road networks. All three models correctly capture this phenomenon (i.e., a higher peak for I_{model} when ρ is higher). However, by not explicitly considering the network topology, both classical and average-degree-based SIR models showed higher node recovery, resulting in disproportionately inflating the number of congested nodes over the considered time period.

Comparing the two cities, despite having different road topology and traffic demand patterns, they have a similar number of congested nodes at peak for both ρ = 0.4 and 0.5 (see Figures 5.7 and 5.8). For a large ρ (ρ = 0.6), the observed difference between the data and the model grows. This is mainly due to the looser definition of congestion when ρ is large, more nodes are considered congested even though they are relatively still flowing at a decent speed. Therefore, we expect to see a divergence between cities as ρ increases.



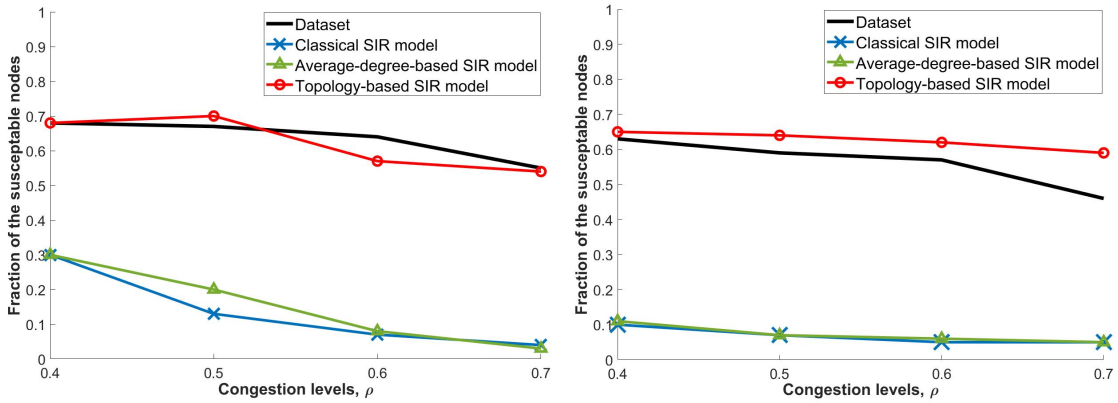
(a) Recovered nodes PEMS-BAY (b) Recovered nodes METR-LA

Figure 5.9: The evolution of S_{LCC} curve (normalised) for different disruption strategies in the PEMS-BAY and METR-LA road networks (Kozhabek et al. 2024).

We show in Figure 5.9 the fraction of nodes recovered from congestion at the end of the observation period across different ρ for PEMS-BAY and METR-LA. In

both networks, we see the topology-based SIR model closely follows the ground truth. The other two models, which have disregarded the topology, have failed to correctly predict the congestion state. Both the classical and average-degree-based SIR models have made similar predictions. The discrepancies with the ground truth are significant. For instance, for PEMS-BAY, the average difference between the actual traffic conditions and that computed by the topology-based SIR model is 0.025.

In contrast, the differences for classical and average-degree-based SIR models are 0.5 and 0.48, respectively. Interestingly, for PEMS-BAY, we see that the model predicted a higher number of recovered nodes despite predicting a lower peak (cf. Figure 5.7). This is because the model computed that the congestion spreads to more nodes but is distributed more evenly over the period; thus, resulting in the peak being not as high but having more congested nodes. Our work further corroborates the literature (e.g., Jiang and Claramunt 2004; Wu et al. 2006; Brockmann and Helbing 2013; Saberi et al. 2020) that traffic congestion spreading mimics the contagion process of an epidemic.



(a) Susceptible nodes PEMS-BAY (b) Susceptible nodes METR-LA

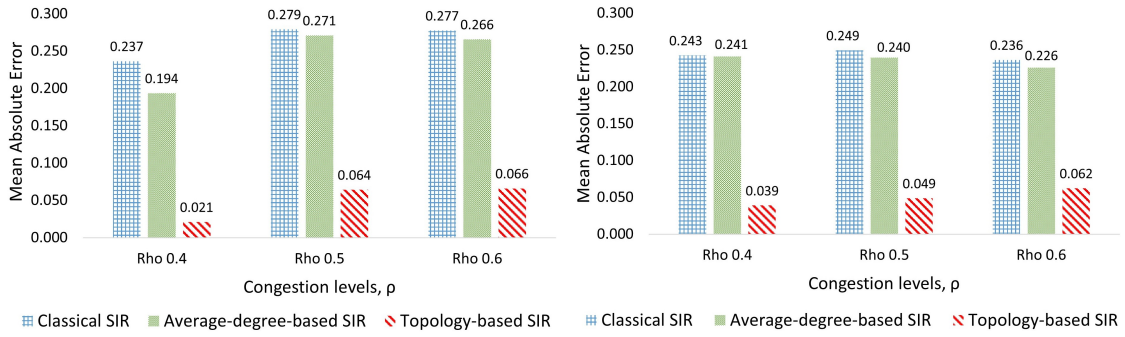
Figure 5.10: (a)-(b) fraction of susceptible nodes from data and models at $\rho = 0.4 - 0.7$ for PEMS-BAY and METR-LA, respectively (Kozhabek et al. 2024).

We correspondingly show in Figure 5.10 the fraction of nodes that never get congested across different ρ for both networks. Similar insights could be found here whereby both classical and average-degree-based SIR models were unable

to track the ground truth while our topology-based SIR model made successful predictions (the average discrepancies are 0.0275 and 0.07 for PEMS-BAY and METR-LA, respectively). In addition, we also note that the discrepancies for our topology-based SIR model are higher in METR-LA than in PEMS-BAY. This is due to the fact that traffic in METR-LA has higher volatility.

5.3.3 Spatio-temporal evolution of congestion

One feature of the topology-based SIR model is the ability to compute the state of each node individually. As such, we can create a congestion map over time to show spatially the spread of traffic congestion. This is not possible for both the classical and average-degree-based SIR models since they treat the network as a whole and compute the aggregated fraction of nodes in each state.

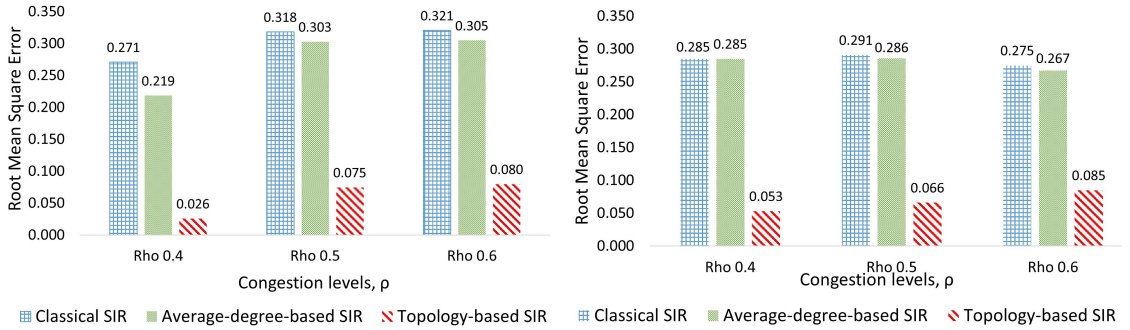


(a) PEMS-BAY (b) METR-LA

Figure 5.11: Mean Absolute Error (MAE) of three models: (a) PEMS-BAY, (b) METR-LA (Kozhabek et al. 2024).

In the following, we present the mean absolute error (MAE) and root mean square error (RMSE) comparing the models against the ground truth in Figure 5.11 and Figure 5.12, respectively. From the figures, we can see that the average-degree-based SIR model marginally performs better than the classical SIR model but our topology-based SIR model achieves significantly lower errors against the other two models.

Finally, we illustrate the spatio-temporal evolution of traffic congestion by showing how the state of the nodes is distributed geographically on the map. We use a



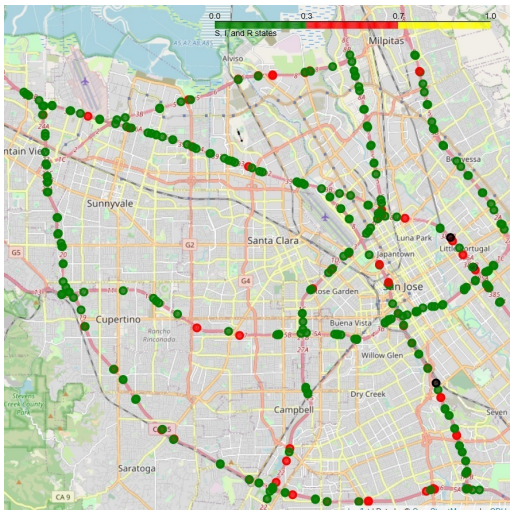
(a) PEMS-BAY (b) METR-LA

Figure 5.12: Root Mean Square Error (RMSE) of three models: (a) PEMS-BAY, (b) METR-LA (Kozhabek et al. 2024).

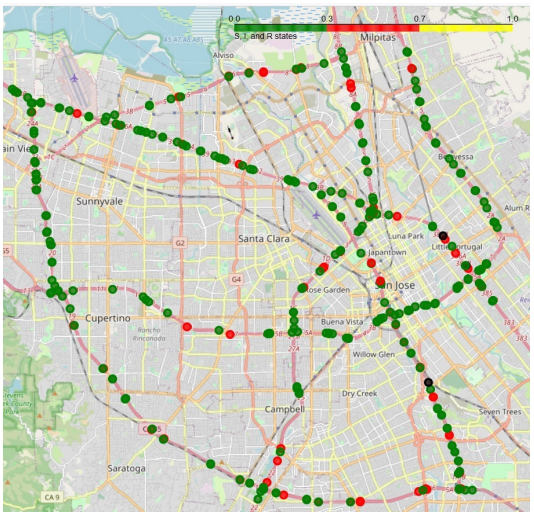
series of hourly congestion maps comparing the actual congestion (derived from the dataset) and that computed by the topology-based SIR model in Figure 5.13 and Figure 5.14 for PEMS-BAY and METR-LA, respectively. For both cases, the initial congested nodes are based on the dataset (seed nodes depicted in black in the maps). We set the same seed nodes for the model. In the figures, nodes in green indicate the free flow traffic state, nodes in red indicate congested locations and nodes in yellow indicate locations that have recovered from the congestion state during the considered period.

In general, we can see that the model offers a good approximation of the conditions of the traffic at different areas of roads. For PEMS-BAY, we can see that at 6 am, the network is largely free flowing but by 7 am, several roads are already congested (e.g., top right corner). Beginning from 8 am, the congestion gradually dissipates (more nodes in yellow appearing). These are all captured by the model (see the corresponding hourly maps Figure 5.13 (b), (d), (f), (h), and (j)).

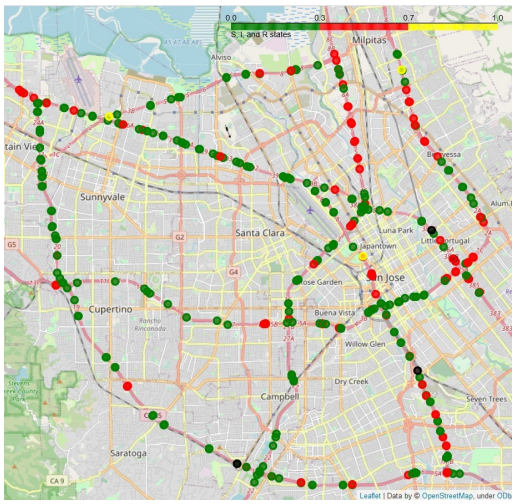
For METR-LA, the data showed a slower build-up of congestion in which the network stayed relatively free-flowing between 6 am - 7 am and congestion only gradually appearing in the later hours. We can see from the counterpart figures computed by the model equally up to the task in capturing this spreading. Therefore, the model can provide a good overall estimation of the traffic conditions in terms of geographical areas but we also note that there are some inaccuracies



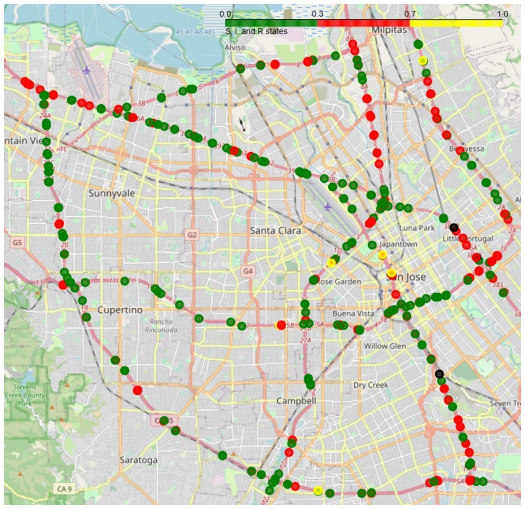
(a) Map from the dataset 06:00 am



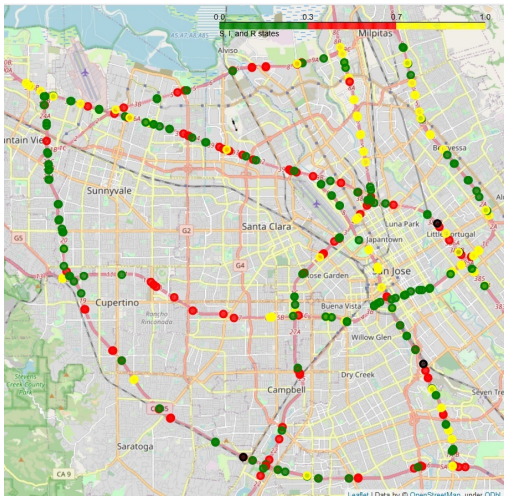
(b) Map based on the model 06:00 am



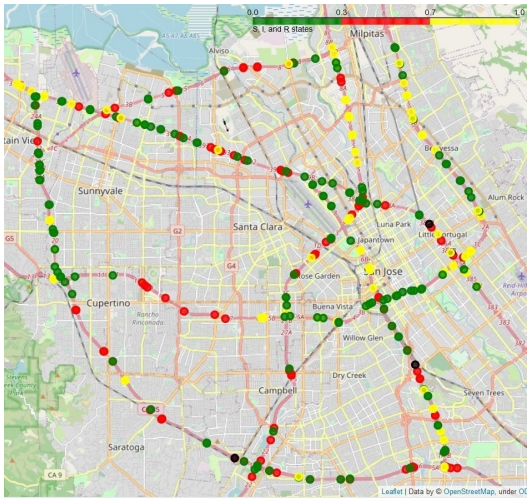
(c) Map from the dataset 07:00 am



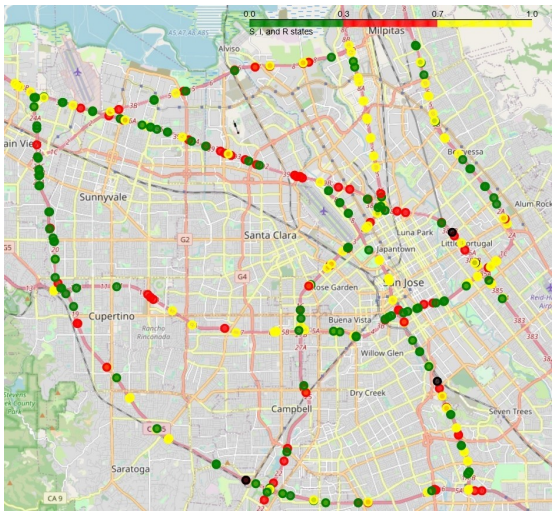
(d) Map based on the model 07:00 am



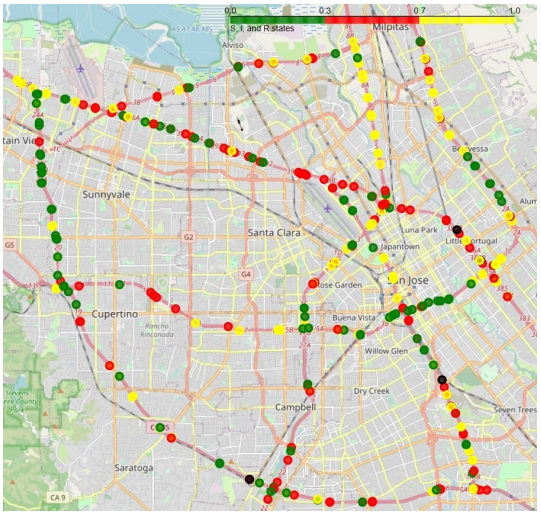
(e) Map from the dataset 08:00 am



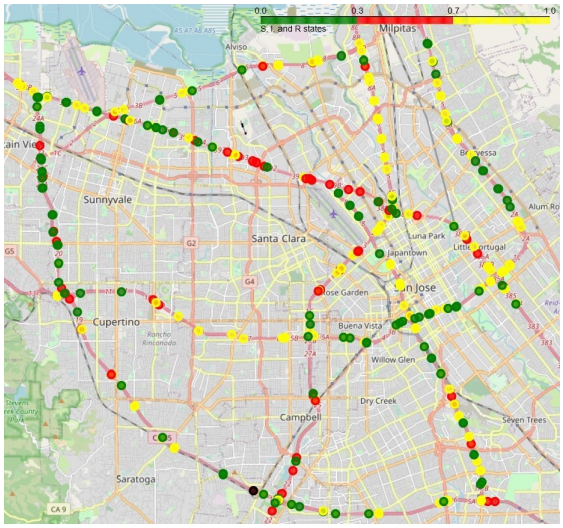
(f) Map based on the model 08:00 am



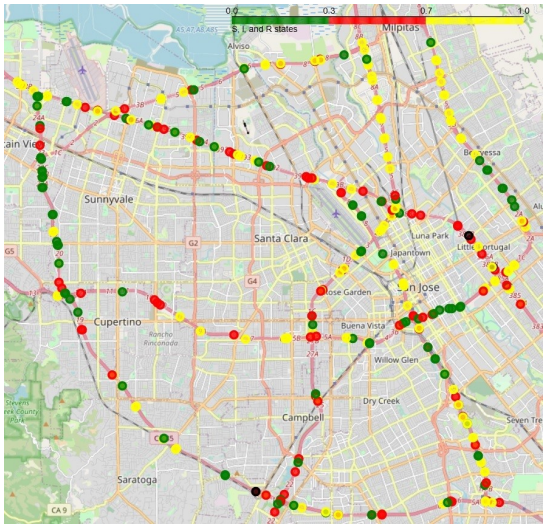
(g) Map from the dataset 09:00 am



(h) Map based on the model 09:00 am

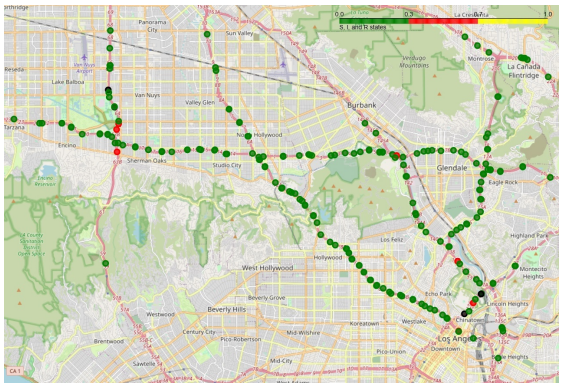


(i) Map from the dataset 10:00 am

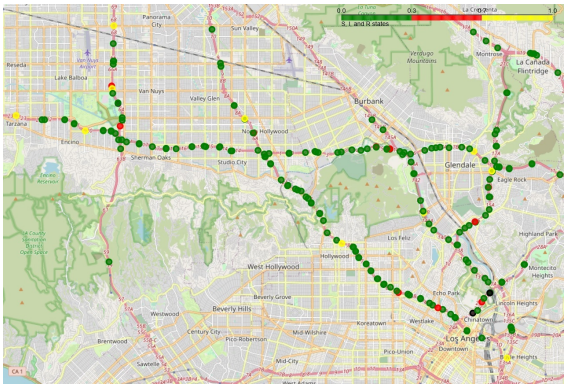


(j) Map based on the model 10:00 am

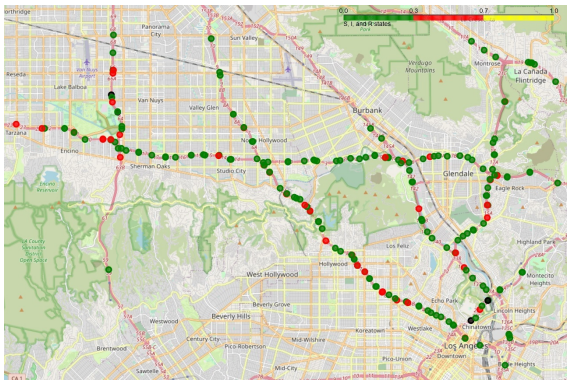
Figure 5.13: Hourly congestion map based on dataset and the model for PEMS-BAY. The first column represents dataset and the second column illustrates output of the model (Kozhabek et al. 2024).



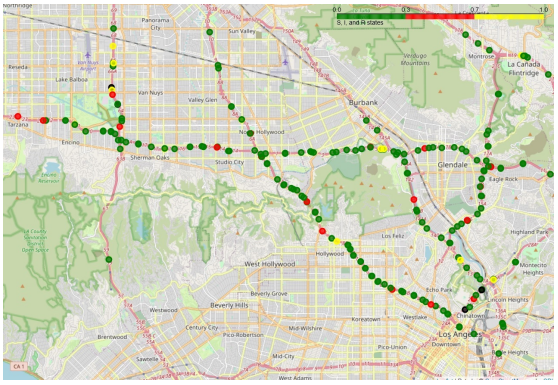
(a) Map from the dataset 06:00 am



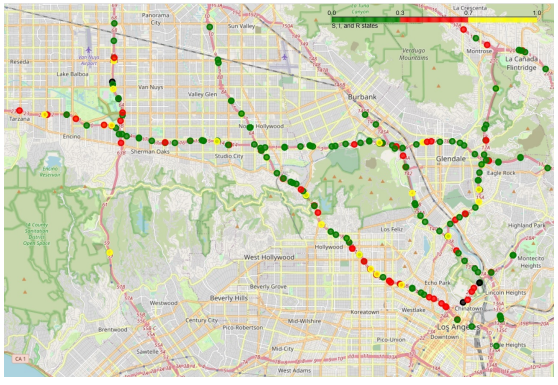
(b) Map based on the model 06:00 am



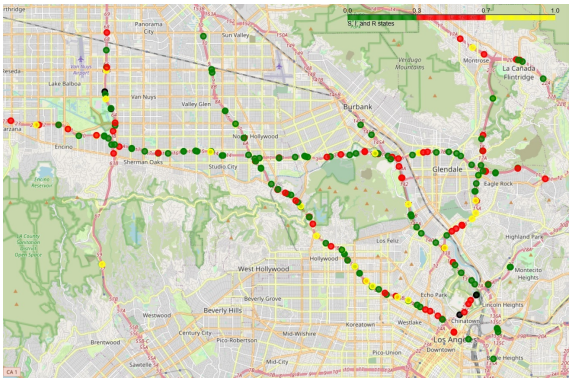
(c) Map from the dataset 07:00 am



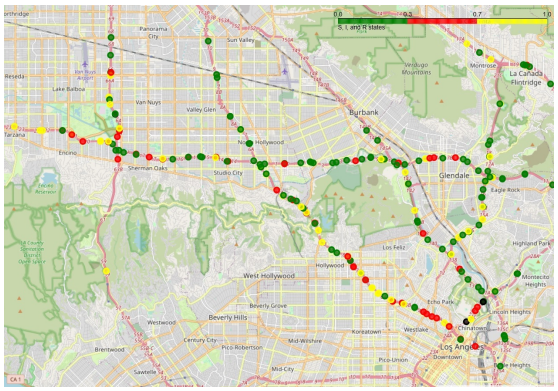
(d) Map based on the model 07:00 am



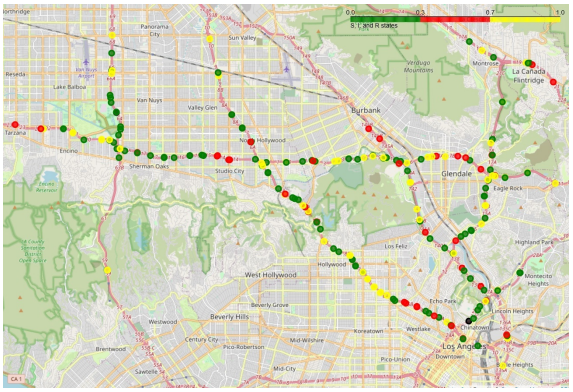
(e) Map from the dataset 08:00 am



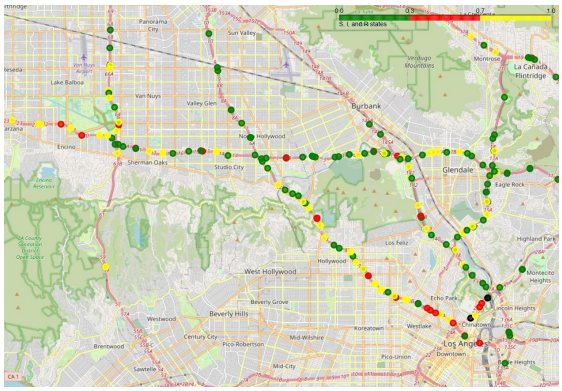
(f) Map based on the model 08:00 am



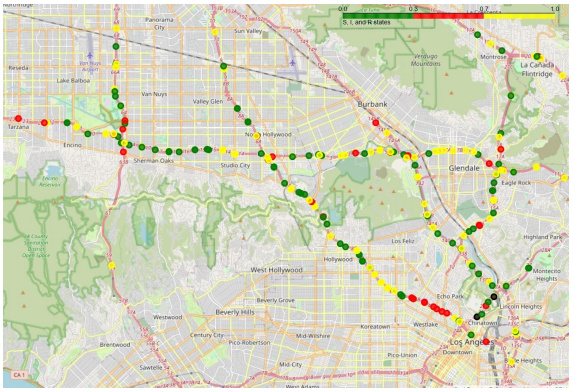
(g) Map from the dataset 09:00 am



(h) Map based on the model 09:00 am



(i) Map from the dataset 10:00 am



(j) Map based on the model 10:00 am

Figure 5.14: Hourly congestion map based on the dataset and the model for METR-LA. The first column represents dataset and the second column illustrates output of the model (Kozhabek et al. 2024).

when specific nodes are of interest.

5.4 Chapter Remarks

In this chapter, we first advocate the use of the epidemic model for modeling the spreading of traffic congestion. Despite the complex nature of urban traffic, we show in this work that the contagion-like process of traffic congestion can be modeled using the SIR epidemic model that includes propagation and dissipation characteristics of traffic dependent on time-varying travel demand. Departing from the classical and average-degree-based SIR models, which mostly consider homogeneous mixing based on the law of mass action (Daley and Gani 2001; Kermack and McKendrick 1927), we propose the explicit consideration of the road network structure via the use of the corresponding adjacency matrix that represents the network topology.

Further, we tested the model against traffic data from two cities, namely California and Los Angeles, and compared the performance of the model against the classical SIR and average-degree-based SIR models. We show that the proposed topology-based SIR model outperforms these two SIR models by a significant margin. Our results also demonstrate better agreement, both temporally and spatially, between the model and data despite the two cities having different traffic profiles. We also illustrate the congestion spreading spatially over time; a feature of our model that the other models are incapable of computing. Our model can be applied to develop control strategies with different objectives, such as minimising the total duration of congestion, minimising the total number of congested roads and junctions, and minimising the recovery time.

Chapter 6

Conclusion and Future Directions

6.1 Summary

The study of urban road networks is crucial for the efficient functioning and development of cities as they serve as the backbone of transportation systems. This thesis provides a comprehensive analysis of both the static and dynamic properties of urban road networks. In this context, static structures represent the consistent topological characteristics and spatial design elements that shape the layout of the network. Evaluating the robustness of a road network against disturbances requires analysing how its configuration and strength endure obstacles like node removal according to established criteria. Conversely, dynamic processes relate to the evolving behaviors over time, traffic dynamics, and shifting patterns occurring within the network. By investigating both static and dynamic elements, this study aims to provide a holistic understanding of urban road networks and their implications for urban planning and management.

The thesis delves into a comprehensive analysis of urban road networks in densely populated cities, utilising tools from network science to explore their topological properties at various scales. At the macro-scale, correlations among different network metrics are identified, revealing insights into the connectivity and efficiency

of urban road networks. The meso-scale analysis uncovers high modularity and polycentric structures within these networks, shedding light on the presence of multiple regions with concentrated road connectivity. Moving to the micro-scale, the study highlights distribution patterns of centrality measures, emphasising the physical constraints influencing the construction of road infrastructures. These multi-scale analyses offer valuable insights for urban planning and development, especially in growing cities.

Furthermore, the thesis investigates the robustness of urban road networks through a node removal process, assessing the impact on network connectivity and efficiency under various disruption strategies. Results indicate the differential effects of random and centrality-based disruption strategies on network degradation, highlighting the importance of targeted perturbations for understanding network vulnerabilities. Recommendations are provided for assessing robustness based on different disruption strategies, emphasising the significance of considering centrality measures that account for network paths.

Another focus of this research was to explore and suggest a straightforward yet effective approach for describing traffic congestion dynamics in extensive networks. The occurrence of traffic jams in urban networks is commonly perceived as a complicated spatial and temporal occurrence. Typically, analysing this phenomenon requires the use of detailed models that demand substantial computational resources and numerous model parameters for calibration. However, this chapter demonstrates that a straightforward contagious process can effectively describe how traffic congestion spreads and diminishes within a network.

Even though traffic congestion is driven by intricate human behavior, it was illustrated that congestion in urban networks follows a remarkably similar spread pattern as seen in other systems, such as the transmission of diseases in a population or the dissemination of ideas in social networks. These phenomena can be elucidated using a concise theoretical network framework. Specifically, the study adapted a classic epidemic model to describe congestion spread in urban networks, considering a propagation and dissipation mechanism that relies on dynamic travel demand and aligns with network traffic flow theory fundamentals.

The research introduced two key features for understanding network traffic dynamics: congestion propagation rate and congestion dissipation rate.

The efficacy and relevance of the proposed contagion-based model were confirmed through empirical analysis of various cities. This model can be employed to observe, predict, and manage the proportion of congested nodes in a network over time, providing potential for adaptive and predictive congestion control in urban networks. By real-time monitoring of the network and tracking congested links, the model could support the development of optimal control strategies with various objectives, such as reducing total congestion duration, minimising the overall number of congested links, or decreasing recovery time. Similar to other control applications based on macroscopic models, the proposed SIR-based model.

Ultimately, the thesis contributes a comprehensive understanding of urban road networks, their topological properties, robustness characteristics, and the dynamics of traffic congestion spreading, providing essential knowledge for urban planners, policymakers, and stakeholders involved in city development and transportation management.

6.2 Future Directions

This section presents several potential directions for future research, offering opportunities for further exploration and development in the field. One key emphasis is the vital necessity of widening the scope of road network comparisons to encompass a more comprehensive understanding of the complex interplay between various factors influencing URNs.

It is believed that the diversity in social, economic, cultural, and geographical factors may impact the development of URNs (Rode et al. 2017), thus potentially molding the structural features of these networks. For example, examining road networks in different geographical contexts, such as coastal urban areas like Sydney, Australia, and mountainous landlocked cities like Almaty, Kazakhstan, could reveal how natural landscapes and terrain affect the design and effectiveness of

urban road networks. Contrasting road networks in culturally diverse cities such as Istanbul, Turkey, and homogenous cities like Oslo, Norway, could demonstrate how cultural preferences and practices impact road design and usage patterns. Further research in this line could reveal distinct topological or operational aspects of URNs based on a geographically diverse selection of test networks.

Another potential future research path could be modeling congestion propagation, considering diverse traffic scenarios that present unique challenges. Examples of such scenarios include evacuation procedures during emergencies, management of traffic flow during mega-events (Elkhouly et al. 2023), routing for VIP movements (Dutta et al. 2023), analysing congestion patterns during low-traffic times, and many more. Exploring these specialized cases offers the opportunity to enhance our understanding of how road networks function under varying conditions and constraints.

For a more comprehensive understanding of traffic patterns under varying conditions, conducting additional research across multiple days (Zang et al. 2023), including weekends, and special events, would be advantageous. By including weekends, and special events in the study, researchers can gather a wide range of data that reflects the varying dynamics of traffic flow. This approach allows for a more thorough analysis of how factors such as daily routines, leisure activities, and exceptional events impact congestion and traffic behavior on road networks. The US Federal Highway Administration (FHWA) outlined main factors contributing to congestion: insufficient highway capacity, poorly maintained traffic control systems, fluctuating demand throughout the day, traffic accidents, road construction projects, and special events such as sports games (Systematics Cambridge 2005; Falcocchio and Levinson 2015). Studying traffic patterns across different days can provide a holistic view of the complexities involved, offering valuable insights.

While this research provides a detailed analysis of the physical condition of the road network using unweighted and undirected data through complex network analysis, exploring the weighted and directed network could offer a more insightful understanding of network dynamics. Additionally, expanding the study to include

the removal of significant edges in a directed network would provide a deeper insight into topological characteristics by pinpointing edges with higher flow rates.

Bibliography

Abdi, A., Saffarzadeh, M. and Salehikalam, A., 2016. Identifying and analyzing stop and go traffic based on asymmetric theory of driving behavior in acceleration and deceleration. *Int. J. Transp. Eng.*, 3 (4), 237–251.

Abdullah, S. and Wu, X., 2011. An epidemic model for news spreading on twitter. *2011 IEEE 23rd Int. Conf. Tools Artif. Intell.*, IEEE, 163–169.

Ahmed, H., Huang, Y. and Lu, P., 2021. A review of car-following models and modeling tools for human and autonomous-ready driving behaviors in microsimulation. *Smart Cities*, 4 (1), 314–335.

Akbarzadeh, M., Memarmontazerin, S. and Soleimani, S., 2018. Where to look for power laws in urban road networks? *Appl. Netw. Sci.*, 3, 1–11.

Akçelik, R., 2007. A review of gap-acceptance capacity models. *Conference of Australian Institutes of Transport Research, Adelaida, Australia*, 5–29.

Albert, R. and Barabási, A., 2002. Statistical mechanics of complex networks. *Rev. Mod. Phys.*, 74 (1), 47.

Albert, R., Jeong, H. and Barabási, A., 2000. Error and attack tolerance of complex networks. *Nature*, 406 (6794), 378–382.

Alrumaih, T. and Alenazi, M., 2023. Evaluation of industrial network robustness against targeted attacks. *Concurrency Comput. Pract. Exp.*.

Anbaroglu, B., Heydecker, B. and Cheng, T., 2014. Spatio-temporal clustering for

non-recurrent traffic congestion detection on urban road networks. *Transp. Res. Part C: Emerg. Technol.*, 48, 47–65.

Anderson, R. and May, R., 1991. *Infectious diseases of humans: dynamics and control*. Oxford UP.

Anvaka GitHub, 2024. An Anvaka city-roads open-source visualization web tool written in JavaScript that extracts data from OSM to draw all the streets within a city. It is powered by WebGL. https://anvaka.github.io/city-roads/. Accessed: 08/03/2024.

Arnott, R., 2013. A bathtub model of downtown traffic congestion. *J. Urban Econ.*, 76, 110–121.

Badhrudeen, M., Derrible, S., Verma, T., Kermanshah, A. and Furno, A., 2022. A geometric classification of world urban road networks. *Urban Sci.*, 6 (1), 11.

Bansal, P. and Graham, D., 2023. Congestion in cities: Can road capacity expansions provide a solution? *Transp. Res. Part A Policy Pract.*, 174, 103726.

Barabasi, A., 2013. Network science. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.*, 371 (1987), 20120375.

Barabasi, A., 2014. Network science network robustness. http://networksciencebook.com/chapter/8. Accessed: 01/03/2024.

Barabási, A., 2016. *Network Science Spreading Phenomena*. Cambridge UP.

Barthélemy, M., 2011. Spatial networks. *Physics Reports*, 499 (1), 1–101.

Barthelemy, M., 2016. *The structure and dynamics of cities*. Cambridge UP.

Barthelemy, M., 2018. *Morphogenesis of spatial networks*. Springer.

Barthelemy, M., Bordin, P., Berestycki, H. and Gribaudi, M., 2013. Self-organization versus top-down planning in the evolution of a city. *Sci. Rep.*, 3 (1), 2153.

Bellingeri, M., Bevacqua, D., Scotognella, F., Alfieri, R. and Cassi, D., 2020. A comparative analysis of link removal strategies in real complex weighted networks. *Sci. Rep.*, 10 (1), 3911.

Ben, S., 2019. Significance of road infrastructure on economic sustainability. *AI-JMSR*, 5 (4), 1–9.

Bettencourt, L., 2013. The origins of scaling in cities. *Science*, 340 (6139), 1438–1441.

Bettencourt, L. and West, G., 2010. A unified theory of urban living. *Nature*, 467 (7318), 912–913.

Blanchard, P., 2009. *Mathematical analysis of urban spatial networks*. Springer: Berlin/Heidelberg. 151–157.

Boccaletti, S., Latora, V., Moreno, Y., Chavez, M. and Hwang, D., 2006. Complex networks: Structure and dynamics. *Physics Reports*, 424 (4-5), 175–308.

Boeing, G., 2017. Osmnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks. *Comput. Environ. Urban Syst.*, 65, 126–139.

Boeing, G., 2018. Measuring the complexity of urban form and design. *Urban Des. Int.*, 23 (4), 281–292.

Boguná, M., Pastor-Satorras, R. and Vespignani, A., 2003. Absence of epidemic threshold in scale-free networks with degree correlations. *Phys. Rev. Lett.*, 90 (2), 028701.

Borgatti, S. and Everett, M., 2000. Models of core/periphery structures. *Soc. Netw.*, 21 (4), 375–395.

Boston Consulting Group, 2018. Barcelona urban lab: using the city as a testing ground for innovation. https://www.centreforpublicimpact.org/case-study/barcelona-urban-lab. Accessed: 07/03/2024.

Brackstone, M. and McDonald, M., 1999. Car-following: a historical review. *Transp. Res. Part F*, 2 (4), 181–196.

Brandes, U., Delling, D., Gaertler, M., Görke, R., Hoefer, M., Nikoloski, Z. and Wagner, D., 2006. On modularity-np-completeness and beyond. *ITI Wagner Tech. Rep.*, 19.

Bretagnolle, A., Daudé, E. and Pumain, D., 2006. From theory to modelling: urban systems as complex systems. *Cybergeo: Eur. J. Geogr.*, 335, 1–17.

Brinkmeier, M. and Schank, T., 2005. Network statistics. *Network analysis: methodological foundation*, Springer, 293–317.

Brockmann, D. and Helbing, D., 2013. The hidden geometry of complex, network-driven contagion phenomena. *Science*, 342 (6164), 1337–1342.

Buhl, J., Gautrais, J., Reeves, N., Solé, R., Valverde, S., Kuntz, P. and Theraulaz, G., 2006. Topological patterns in street networks of self-organized urban settlements. *Eur. Phys. J. B.*, 49, 513.

Burghardt, K., Uhl, J., Lerman, K. and Leyk, S., 2022. Road network evolution in the urban and rural US since 1900. *Comput. Environ. Urban Syst.*, 95, 101803.

Cardillo, A., Scellato, S., Latora, V. and Porta, S., 2006. Structural properties of planar graphs of urban street patterns. *Phys. Rev. E*, 73 (6), 066107.

Carlton, R., 2023. Forbes: 7 Million U.K. Drivers Plan To Cut Back Car Use For Short Journeys, Finds Kwik Fit. https://www.forbes.com/sites/carltonreid/2023/04/24/7-million-uk-drivers-plan-to-cut-back-car-use-for-short-journeys-finds-kwik-fit/. Accessed: 02/03/2024.

Casali, Y. and Heinimann, H., 2019. A topological analysis of growth in the Zurich road network. *Comput. Environ. Urban Syst.*, 75, 244–253.

Casali, Y. and Heinimann, H., 2020. Robustness response of the Zurich road network under different disruption processes. *Comput. Environ. Urban Syst.*, 81, 101460.

Chai, W., 2017. Modelling spreading process induced by agent mobility in complex networks. *IEEE Trans. Netw. Sci. Eng.*, 5 (4), 336–349.

Chai, W., He, D., Psaras, I. and Pavlou, G., 2013. Cache “less for more” in information-centric networks. *Comput. Commun.*, 36 (7), 758–770.

Chai, W., Kyritsis, V., Katsaros, K. and Pavlou, G., 2016. Resilience of interdependent communication and power distribution networks against cascading failures. *IFIP Networking Conference*, IEEE, 37–45.

Chai, W. and Pavlou, G., 2017. Path-based epidemic spreading in networks. *IEEE/ACM Trans. Netw.*, 25 (1), 565–578.

Chehreghani, M. and Chehreghani, M., 2014. Modeling transitivity in complex networks. *arXiv preprint:1411.0958*.

Chen, C., 2002. *Freeway performance measurement system (PeMS)*. University of California, Berkeley.

Chen, Y., Mao, J., Zhang, Z., Huang, H., Lu, W., Yan, Q. and Liu, L., 2022. A quasi-contagion process modeling and characteristic analysis for real-world urban traffic network congestion patterns. *Phys. A: Stat. Mech. Appl.*, 603, 127729.

Chinese Ministry of Transport, 2014. Guidelines for the construction of urban public transportation smart application demonstration projects.

Clauset, A., Newman, M. and Moore, C., 2004. Finding community structure in very large networks. *Phys. Rev. E*, 70 (6), 066111.

Colizza, V., Barrat, A., Barthelemy, M., Valleron, A. and Vespignani, A., 2007. Modeling the worldwide spread of pandemic influenza: baseline case and containment interventions. *PLoS medicine*, 4 (1), e13.

Costa, L., Travençolo, B., Viana, M. and Strano, E., 2010. On the efficiency of transportation systems in large cities. *Eur. Phys. Lett.*, 91 (1), 18003.

Craig, T., Hermann, P., Cox, E. and Weiner, R., 2024. Baltimore bridge collapse and port closure 'devastating,' residents say. https://www.washingtonpost.com/dc-md-va/2024/03/26/baltimore-bridge-collapse-impact-traffic-jobs/. Accessed: 27/03/2024.

Crucitti, P., Latora, V. and Porta, S., 2006. Centrality in networks of urban streets. *Chaos: Int. J. Nonlinear Sci.*, 16 (15113).

Cucuringu, M., Rombach, P., Lee, S. and Porter, M., 2016. Detection of core-periphery structure in networks using spectral methods and geodesic paths. *Eur. J. Appl. Math.*, 27 (6), 846–887.

Daley, D. and Gani, J., 2001. *Epidemic modelling: an introduction*, volume 15. Cambridge UP. 20–37.

Derrible, S. and Kennedy, C., 2010. The complexity and robustness of metro networks. *Phys. A: Stat. Mech. Appl.*, 389 (17), 3678–3691.

Diehl, F., Brunner, T., Le, M. and Knoll, A., 2019. Graph neural networks for modelling traffic participant interaction. *IEEE Intell. Veh. Symp. (IV)*, 695–701.

Diestel, R., 2024. *Graph theory*. Springer.

Diop, I., Cherifi, C., Diallo, C. and Cherifi, H., 2021. Revealing the component structure of the world air transportation network. *Appl. Netw. Sci.*, 6, 1–50.

Diop, I., Cherifi, C., Diallo, C. and Cherifi, H., 2022. Robustness of the weighted world air transportation network components. *IEEE COMPENG*, IEEE, 1–6.

Duan, Y. and Lu, F., 2013. Structural robustness of city road networks based on community. *Comput. Environ. Urban Syst.*, 41, 75–87.

Duan, Y. and Lu, F., 2014. Robustness of city road networks at different granularities. *Phys. A: Stat. Mech. Appl.*, 411, 21–34.

Dutta, P., Khatua, S. and Choudhury, S., 2023. Fast move: A prioritized vehicle rerouting strategy in smart city. *Veh. Comm.*, 44, 100666.

Elkhouly, R., Tamaki, E. and Iwasaki, K., 2023. Mitigating crowded transportation terminals nearby mega-sports events. *Behav. Inf. Technol.*, 42 (7), 904–920.

Erdos, P. and Renyi, A., 1959. On random graphs. *Publ. Math.*, 18.

Falcocchio, J. C. and Levinson, H. S., 2015. *Road traffic congestion: a concise guide*, volume 7. Springer.

Faloutsos, M., Faloutsos, P. and Faloutsos, C., 1999. On power-law relationships of the internet topology. *ACM CCS*, 29 (4), 251–262.

Fan, Z. and Harper, C., 2022. Congestion and environmental impacts of short car trip replacement with micromobility modes. *Transp. Res. D Trans. Environ*, 103, 103173.

Fu, R., Zhang, Z. and Li, L., 2016. Using lstm and gru neural network methods for traffic flow prediction. *31st YAA Conference, Chinese Assoc. of Automation*, IEEE, 324–328.

Glaeser, E., 2011. Cities, productivity, and quality of life. *Science*, 333 (6042), 592–594.

Goh, K.-I., Kahng, B. and Kim, D., 2001. Universal behavior of load distribution in scale-free networks. *Phys. Rev. Lett.*, 87 (27), 278701.

Goh, S., Choi, M., Lee, K. and Kim, K., 2016. How complexity emerges in urban systems: Theory of urban morphology. *Phys. Rev. E*, 93 (5), 052309.

H., L., X., C. and J, L., 2014. Understanding video sharing propagation in social networks: Measurement and analysis. *ACM Trans. Multimedia Comput. Commun. Appl.*, 10 (4), 1–20.

Habitat, U., 2022. Envisaging the future of cities. *World Cities Report*.

Helbing, D., 1996. Gas-kinetic derivation of navier-stokes-like traffic equations. *Phys. Rev. E*, 53 (3), 2366.

Hethcote, H. and Yorke, J., 1984. Lecture notes in biomathematics.

Hopkins, B. and Wilson, R., 2004. The truth about königsberg. *Coll. Math. J.*, 35 (3), 198–207.

Huang, G., 2014. An insight into extreme learning machines: random neurons, random features and kernels. *Cogn. Comput.*, 6, 376–390.

Immers, B., Yperman, I., Stada, J. and Bleukx, A., 2004. Reliability and robustness of transportation networks; problem survey and examples. *Proc. of the NECTAR Clust. Meeting on Reliab. of Networks*, 1–12.

Iturria-Medina, Y., Sotero, R., Toussaint, P., Evans, A. and Initiative, A. D. N., 2014. Epidemic spreading model to characterize misfolded proteins propagation in aging and associated neurodegenerative disorders. *PLoS*, 10 (11), e1003956.

Iyer, S., Killingback, T., Sundaram, B. and Wang, Z., 2013. Attack robustness and centrality of complex networks. *PloS one*, 8 (4), e59613.

Jagadish, H., Gehrke, J., Labrinidis, A., Papakonstantinou, Y., Patel, J., Ramakrishnan, R. and Shahabi, C., 2014. Big data and its technical challenges. *Commun. ACM*, 57 (7), 86–94.

Jiang, B. and Claramunt, C., 2004. Topological analysis of urban street networks. *Environ. Plann. B: Plann. Des.*, 31 (1), 151–162.

Jiang, B., Zhao, S. and Yin, J., 2008. Self-organized natural roads for predicting traffic flow: a sensitivity study. *J. Stat. Mech. Theory Exp.*, (07), P07008.

Justen, A., Martinez, F. and Cortes, C., 2013. The use of space–time constraints for the selection of discretionary activity locations. *J. Transp. Geogr.*, 33, 146–152.

Kalapala, V., Sanwalani, V., Clauset, A. and Moore, C., 2006. Scale invariance in road networks. *Phys. Rev. E*, 73 (2), 026130.

Karduni, A., Kermanshah, A. and Derrible, S., 2016. A protocol to convert spatial polyline data to network formats and applications to world urban road networks. *Scientific Data*, 3 (1), 1–7.

Karlberg, M., 1997. Testing transitivity in graphs. *Soc. Netw.*, 19 (4), 325–343.

Keeling, M. and Eames, K., 2005. Networks and epidemic models. *J. R. Soc. Interface*, 2 (4), 295–307.

Kermack, W. and McKendrick, A., 1927. A contribution to the mathematical theory of epidemics. *Proc. R. Soc. Lond. Ser. A-Contain. Pap. Math. Phys.*, 115 (772), 700–721.

Kirkley, A., Barbosa, H., Barthelemy, M. and Ghoshal, G., 2018. From the betweenness centrality in street networks to structural invariants in random planar graphs. *Nature*, 9 (1), 2501.

Kloosterman, R. and Musterd, S., 2001. The polycentric urban region: towards a research agenda. *Urban Studies*, 38 (4), 623–633.

Knaap, E. and Rey, S., 2023. Segregated by design? street network topological structure and the measurement of urban segregation. *Environ. Plan. B: Urban Analytics City Sci*.

Ko, E., Ahn, J. and Kim, E., 2016. 3d markov process for traffic flow prediction in real-time. *Sensors*, 16 (2), 147.

Kojaku, S. and Masuda, N., 2017. Finding multiple core-periphery pairs in networks. *Phys. Rev. E*, 96 (5), 052313.

Koulakezian, A., Abdelgawad, H., Tizghadam, A., Abdulhai, B. and Leon-Garcia, A., 2015. Robust network design for roadway networks: Unifying framework and application. *IEEE Intell. Transp. Syst. Mag.*, 7 (2), 34–46.

Kozhabek, A., Chai, W. K. and Zheng, G., 2024. Modeling traffic congestion spreading using a topology-based SIR epidemic model. *IEEE Access*, 12, 35813–35826.

Kumar, R. and Singh, A., 2020. Robustness in multilayer networks under strategical and random attacks. *Procedia Comput. Sci.*, 173, 94–103.

Kurant, M. and Thiran, P., 2006. Layered complex networks. *Phys. Rev. Lett.*, 96 (13), 138701.

Lämmer, S., Gehlsen, B. and Helbing, D., 2006. Scaling laws in the spatial structure of urban road networks. *Phys. A: Stat. Mech. Appl.*, 363 (1), 89–95.

Lan, T., Zhang, H. and Li, Z., 2022. Exploring the evolution of road centrality: A case study of Hong Kong from 1976 to 2018. *Appl. Geogr*.

Lang, A., N.and Herrmann, World Economic Forum 2022. Micromobility is clean and quiet — how can it be widely used?

Latora, V. and Marchiori, M., 2001. Efficient behavior of small-world networks. *Phys. Rev. Lett.*, 87 (19), 198701.

Latora, V. and Marchiori, M., 2002. Is the boston subway a small-world network? *Phys. A: Stat. Mech. Appl.*, 314 (1-4), 109–113.

Laval, J. and Leclercq, L., 2010. A mechanism to describe the formation and propagation of stop-and-go waves in congested freeway traffic. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.*, 368 (1928), 4519–4541.

Lee, B. and Jung, W., 2018. Analysis on the urban street network of korea: Connections between topology and meta-information. *Phys. A: Stat. Mech. Appl.*, 497, 15–25.

Lee, S., Cucuringu, M. and Porter, M., 2014. Density-based and transport-based core-periphery structures in networks. *Phys. Rev. E*, 89 (3).

Levinson, D., 2012. Network structure and city size. *PloS one*, 7 (1).

Levinson, D. and El-Geneidy, A., 2009. The minimum circuity frontier and the journey to work. *Reg. Sci. Urban Econ.s*, 39 (6), 732–738.

Li, C., Yue, W., Mao, G. and Xu, Z., 2020a. Congestion propagation based bottleneck identification in urban road networks. *IEEE Trans. Veh. Technol.*, 69 (5), 4827–4841.

Li, D., Fu, B., Wang, Y., Lu, G., Berezin, Y., Stanley, H. and Havlin, S., 2015. Percolation transition in dynamical traffic network with evolving critical bottlenecks. *Proc. Natl. Acad. Sci.*, 112 (3), 669–672.

Li, J., Zheng, P. and Zhang, W., 2020b. Identifying the spatial distribution of public transportation trips by node and community characteristics. *Transportation Planning and Technology*, 43 (3), 325–340.

Li, Y., Yu, R., Shahabi, C. and Liu, Y., 2017a. Diffusion convolutional recurrent neural network: Data-driven traffic forecasting. *arXiv preprint:1707.01926*.

Li, Y., Zhao, L., Yu, Z. and Wang, S., 2017b. Traffic flow prediction with big data: A learning approach based on sis-complex networks. *2017 IEEE 2nd ITNEC*, IEEE, 550–554.

Lighthill, M. and Whitham, G., 1955. On kinematic waves ii. a theory of traffic flow on long crowded roads. *Proc. R. Soc. Lond.*, 229 (1178), 317–345.

Lin, J. and Ban, Y., 2013. Complex network topology of transportation systems. *Transport Review*, 33 (6), 658–685.

Lin, J. and Ban, Y., 2017. Comparative analysis on topological structures of urban street networks. *ISPRS International Journal of Geo-Information*, 6 (10), 295.

Liu, W., Li, X., Liu, T. and Liu, B., 2019. Approximating betweenness centrality to identify key nodes in a weighted urban complex transportation network. *J. Adv. Transp.*.

Louf, R. and Barthelemy, M., 2014. A typology of street patterns. *J. R. Soc. Interface*, 11 (101), 20140924.

Ma, F., Shi, W., Yuen, K., Sun, Q., Xu, X., Wang, Y. and Wang, Z., 2020. Exploring the robustness of public transportation for sustainable cities: A double-layered network perspective. *J. Clean. Prod.*, 265, 121747.

Mahmassani, H., Saberi, M. and Zockaie, A., 2013. Urban network gridlock: Theory, characteristics, and dynamics. *Procedia - Soc. Behav. Sci.*, 80, 79–98.

Manzano, M., Marzo, J., Calle, E. and Manolovay, A., 2012. Robustness analysis of real network topologies under multiple failure scenarios. *17th EuCNC*, IEEE, 1–6.

Marinov, T., Marinova, R., J., O. and Jackson, M., 2014. Inverse problem for coefficient identification in sir epidemic models. *Comput. Math. Appl.*, 67 (12), 2218–2227.

Mayor of London, 2022. Cost of congestion in capital revealed as car use remains high. https://www.london.gov.uk/press-releases/mayoral/cost-of-congestion-in-capital-revealed. Accessed: 09/03/2024.

Meloni, S., Arenas, A. and Moreno, Y., 2009. Traffic-driven epidemic spreading in finite-size scale-free networks. *Proc. Natl. Acad. Sci.*, 106 (40), 16897–16902.

Mieghem, P. V., 2011. The n-intertwined sis epidemic network model. *Computing*, 93 (2-4), 147–169.

Mieghem, P. V., Omic, J. and Kooij, R., 2008. Virus spread in networks. *IEEE/ACM Trans. Netw.*, 17 (1), 1–14.

Ming, L., Guohua, S. and Ying, C., 2014. Research on excessive short distance car trips in urban area. *J. Beijing Jiaotong Univ.*, 38 (3), 15–21.

Murray, A. G., 2006. A model of the emergence of infectious pancreatic necrosis virus in scottish salmon farms 1996–2003. *Ecol. Model.*, 199 (1), 64–72.

Nagy, A. and Simon, V., 2021. Traffic congestion propagation identification method in smart cities. *Infocommun. J.*, 13 (1), 45–57.

National Geographic, 2024. Key components of civilization. https://education.nationalgeographic.org/resource/key-components-civilization/. Accessed: 09/03/2024.

Newell, G., 1993. A simplified theory of kinematic waves in highway traffic, part ii: Queueing at freeway bottlenecks. *Transp. Res. D Trans. Environ.*, 27 (4), 289–303.

Newman, M., 2004. Finding and evaluating community structure in networks. *Phys. Rev. E*, 69 (26113), 1–16.

Newman, M., 2012. Networks: An introduction. *Artificial Life*, 18, 241.

Newman, M., Watts, D. and Strogatz, S., 2002. Random graph models of social networks. *PNAS*, 99, 2566–2572.

Nika, M., Wilding, T., Fiems, D., De Turck, K. and Knottenbelt, W., 2015. Going multi-viral: synthedemic modelling of internet-based spreading phenomena. *EAI Trans. Ambient Syst.*, 2 (7).

Novak, D. and Sullivan, J., 2014. A link-focused methodology for evaluating accessibility to emergency services. *Decis. Support Syst.*, 57, 309–319.

Olmos, L. and Muñoz, J., 2017. Traffic gridlock on a honeycomb city. *Phys. Rev. E*, 95 (3), 032320.

Opuszko, M. and Ruhland, J., 2013. Impact of the network structure on the sir model spreading phenomena in online networks. *ICCGI 2013*, 22–28.

Pan, W., Ghoshal, G., Krumme, C., Cebrian, M. and Pentland, A., 2013. Urban characteristics attributable to density-driven tie formation. *Nature Comm.*, 4 (1), 1961.

Parthasarathy, K., Deka, P., Saravanan, S., Abijith, D. and Jennifer, J., 2021. Assessing the impact of 2018 tropical rainfall and the consecutive flood-related damages for the state of kerala, india. *Disaster Resil. Sustain.*, Elsevier, 379–395.

Pastor-Satorras, R. and Vespignani, A., 2001. Epidemic spreading in scale-free networks. *Phys. Rev. Lett.*, 86 (14), 3200.

Pastor-Satorras, R. and Vespignani, A., 2004. *Evolution and structure of the Internet: A statistical physics approach*. Cambridge UP.

Payne, H., 1971. Models of freeway traffic and control. *Math. Models Public Syst.*, 1, 51.

Piccoli, B. and Garavello, M., 2006. Traffic flow on networks. *AIMS*.

Pipes, L., 1953. An operational analysis of traffic dynamics. *J. Appl. Phys.*, 24 (3), 274–281.

Porta, S., Crucitti, P. and Latora, V., 2006. The network analysis of urban streets: a primal approach. *Environ. Plann. B: Plann. Des.*, 33, 705.

Porta, S., Strano, E., Iacoviello, V., Messora, R., Latora, V., Cardillo, A., Wang, F. and Scellato, S., 2009. Street centrality and densities of retail and services in bologna, italy. *Environ. Plann. B: Plann. Des.*, 36 (3), 450–465.

Rahman, M., Chowdhury, M., Xie, Y. and He, Y., 2013. Review of microscopic lane-changing models and future research opportunities. *IEEE Trans. Intell. Transp. Syst.*, 14 (4), 1942–1956.

Reza, S., Machado, J. and Tavares, J., 2022. Analysis of the structure of the road networks: A network science perspective. *ISAIM*, 45–53.

Richards, P., 1956. Shock waves on the highway. *Oper. Res.*, 4 (1), 42–51.

Rode, P., Floater, G., Thomopoulos, N., Docherty, J., Schwinger, P., Mahendra, A. and Fang, W., 2017. Accessibility in cities: transport and urban form. *Disrupt. Mobil.*, 239–273.

Rohe, K., 2013. The difference between the transitivity ratio and the clustering coefficient. https://pages.stat.wisc.edu/ karlrohe/netsci/MeasuringTrianglesInGraphs.pdf. Accessed: 27/02/2024.

Rui, X., Meng, F., Wang, Z., Yuan, G. and Du, C., 2018. Spir: The potential spreaders involved sir model for information diffusion in social networks. *Phys. A: Stat. Mech. Appl.*, 506, 254–269.

Saberi, M., Hamedmoghadam, H., Ashfaq, M., Hosseini, S., Gu, Z., Shafiei, S., Nair, D., Dixit, V., Gardner, L., Waller, S. and González, M., 2020. A simple contagion process describes spreading of traffic jams in urban networks. *Nature Comm.*, 11 (1), 1616.

Sahneh, F., Scoglio, C. and Mieghem, P. V., 2013. Generalized epidemic mean-field model for spreading processes over multilayer complex networks. *IEEE/ACM Trans. Netw.*, 21 (5), 1609–1620.

Saifuzzaman, M. and Zheng, Z., 2014. Incorporating human-factors in car-following models: a review of recent developments and research needs. *Transp. Res. Part C: Emerg. Technol.*, 48, 379–403.

Sakakibara, H., Kajitani, Y. and Okada, N., 2004. Road network robustness for avoiding functional isolation in disasters. *J. Transp. Eng.*, 130 (5), 560–567.

Sanderson, D., Busquets, D. and Pitt, J., 2012. A micro-meso-macro approach to intelligent transportation systems. *IEEE 6th IC-SASO Workshops*, 77–82.

Scardoni, G. and Laudanna, C., 2012. Centralities based analysis of complex networks. *New Front. Graph Theory*, 323–348.

Schank, T. and Wagner, D., 2005. Approximating clustering coefficient and transitivity. *J. Graph Algorithms Appl.*, 9 (2), 265–275.

Schillo, M., Bürckert, H., Fischer, K. and Klusch, M., 2001. Towards a definition of robustness for market-style open multi-agent systems. *Proc. 5th Int. Conf. Autonomous Agents*, 75–76.

Scott, D., Novak, D., Aultman-Hall, L. and Guo, F., 2006. Network robustness index: A new method for identifying critical links and evaluating the performance of transportation networks. *J. Transp. Geogr.*, 14 (3), 215–227.

Shang, W., Chen, Y., Bi, H., Zhang, H., Ma, C. and Ochieng, W., 2020a. Statistical characteristics and community analysis of urban road networks. *Complex.*, 1–21.

Shang, W., Chen, Y., Song, C. and Ochieng, W., 2020b. Robustness analysis of urban road networks from topological and operational perspectives. *Math. Prob. Eng.*, 1.

Sharifi, A., 2019. Resilient urban forms: A review of literature on streets and street networks. *Build. Environ.*, 147, 171–187.

Shulgin, B., Stone, L. and Agur, Z., 1998. Pulse vaccination strategy in the sir epidemic model. *Bulletin of mathematical biology*, 60 (6), 1123–1148.

Siozos-Rousoulis, L., Robert, D. and Verbeke, W., 2021. A study of the us domestic air transportation network: temporal evolution of network topology and robustness from 2001 to 2016. *J. Transp. Secur.*, 14, 55–78.

Smith, P., Hutchison, D., Sterbenz, J., Schöller, M., Fessi, A., Karaliopoulos, M., Lac, C. and Plattner, B., 2011. Network resilience: a systematic approach. *IEEE Commun. Mag.*, 49 (7), 88–97.

Song, Q. and Wang, X., 2010. Efficient routing on large road networks using hierarchical communities. *IEEE Trans. Netw. Sci. Eng.*, 12 (1), 132–140.

Song, Z., Duan, H., Ge, Y. and Qiu, X., 2015. A novel measure of centrality based on betweenness. *CAC*, IEEE, 174–178.

Stanway, D., 2021. Zhengzhou floods serve China's urban planners deadly warning. https://www.reuters.com/world/china/zhengzhou-floods-serve-chinas-urban-planners-deadly-warning-2021-07-23/. Accessed: 18/03/2024.

Strano, E., Strano, E., Nicosia, V., Latora, V., Porta, S. and Barthélemy, M., 2012. Elementary processes governing the evolution of road networks. *Sci. Rep.*, 2 (1), 296.

Strogatz, S., 2001. Exploring complex networks. *Nature*, 410 (6825), 268–276.

Sugiyama, T., Merom, D., van der Ploeg, H., Corpuz, G., Bauman, A. and Owen, N., 2012. Prolonged sitting in cars: prevalence, socio-demographic variations, and trends. *Prev. Med.*, 55 (4), 315–318.

Sun, L., Ling, X., He, K. and Tan, Q., 2016. Community structure in traffic zones based on travel demand. *Phys. A: Stat. Mech. Appl.*, 457, 356–363.

Systematics Cambridge, 2005. Traffic congestion and reliability: Trends and advanced strategies for congestion mitigation. Technical report, U.S. Fed. Highway Admin.

Taipei City Traffic Engineering Office, 2019. Traffic lights get smarter and driving gets more efficient.

Tan, J., Elbaz, K., Wang, Z., Shen, J. and Chen, J., 2020. Lessons learnt from bridge collapse: A view of sustainable management. *Sustain.*, 12 (3), 1205.

Tang, J., Wang, Y. and Liu, F., 2013. Characterizing traffic time series based on complex network theory. *Phys. A: Stat. Mech. Appl.*, 392, 4192.

The Economic Times, 2024. Bharat Bandh today: Routes to avoid as traffic jams expected on major roads in Delhi-NCR. https://economictimes.indiatimes.com/news/india/gramin-bharat-bandh-routes-to-avoid-as-traffic-jams-expected-on-major-roads-in-delhi-ncr-on-feb-16/articleshow/107719431.cms?from=mdr. Accessed: 23/02/2024.

The Times Israel, 2022. Large sinkhole opens on main Tel Aviv highway, causing traffic chaos but no injuries. https://www.timesofisrael.com/large-sinkhole-opens-on-main-tel-aviv-highway-no-injuries/text=A

Toncharoen, R. and Piantanakulchai, M., 2018. Traffic state prediction using convolutional neural network. *15th International Joint Conference on Computer Science and Software Engineering (JCSSE)*, IEEE, 1–6.

Trajanovski, S., Martín-Hernández, J., Winterbach, W. and Van Mieghem, P., 2013. Robustness envelopes of networks. *J. Complex Networks*, 1 (1), 44–62.

Transport for London, 2012. Who travels by car in London and for what purpose? https://content.tfl.gov.uk/technical-note-14-who-travels-by-car-in-london.pdf. Accessed: 03/03/2024.

Transport for London, 2023. Congestion charge marks 20 years of keeping london moving sustainably. https://tfl.gov.uk/info-for/media/press-releases/2023/february/congestion-charge-marks-20-years-of-keeping-london-moving-sustainably. Accessed: 03/03/2024.

Transport Scotland, 2022. A route map to achieve a 20 percent reduction in car kilometers by 2030. https://www.transport.gov.scot/media/50872/a-route-

map-to-achieve-a-20-per-cent-reduction-in-car-kms-by-2030.pdf. Accessed: 02/03/2024.

Tsiotas, D. and Polyzos, S., 2017. The topology of urban road networks and its role to urban mobility. *Transportation Research Procedia*, 24, 482.

UK Department for Transport, 2022. National road traffic projections. https://assets.publishing.service.gov.uk/media/6698c4f90808eaf43b50d193/national-road-traffic-projections-2022.pdf. Accessed: 17/01/2024.

UN DESA, 2019. UN World Urbanization Prospects 2018. https://population.un.org/wup/Publications/Files/WUP2018-Highlights.pdf. Accessed: 27/02/2024.

UN ESCAP, 2021. Evaluation of the project asia-pacific sustainable urban transport systems to support the regions. *Project evaluation report*. Accessed: 30/01/2024.

Vaca-Ramírez, F., 2019. Robustness of urban road networks based on spatial topological patterns. *arXiv preprint*.

Van Mieghem, P., Doerr, C., Wang, H., Hernandez, J., Hutchison, D., Karaliopoulos, M. and Kooij, R., 2010. A framework for computing topological network robustness. *Delft Uni of Tech., Rep. 2010*, 1–15.

Van Tilburg, C., 2007. *Traffic and congestion in the Roman Empire*. Routledge Lond. 42-44.

Vandaele, N., Woensel, T. V. and Verbruggen, A., 2000. A queueing-based traffic flow model. *Transp. Res. D Trans. Environ.*, 5 (2), 121–135.

Verbavatz, V. and Barthelemy, M., 2020. The growth equation of cities. *Nature*, 587 (7834), 397–401.

Viana, M., Strano, E., Bordin, P. and Barthelemy, M., 2013. The simplicity of planar networks. *Sci. Rep.*, 3 (1).

Vickrey, W., 1969. Congestion theory and transport investment. *The American Economic Review*, 251–260.

Wang, F., Antipova, A. and Porta, S., 2011. Street centrality and land use intensity in baton rouge, louisiana. *J. Transp. Geogr.*, 19 (2), 285.

Wang, H., Hernandez, J. and Van Mieghem, P., 2008. Betweenness centrality in a weighted network. *Phys. Rev. E*, 77 (4), 046105.

Wang, P., Hunter, T., Bayen, A., Schechtner, K. and González, M., 2012. Understanding road usage patterns in urban areas. *Sci. Rep.*, 2 (1), 1001.

Wang, X., Koç, Y., Derrible, S., Ahmad, S., Pino, W. and Kooij, R., 2017. Multi-criteria robustness analysis of metro networks. *Phys. A: Stat. Mech. Appl.*, 474, 19–31.

Wang, X., Wu, X., Abdel-Aty, M. and Tremont, P., 2013. Investigation of road network features and safety performance. *Accid. Anal. Prev.*, 56, 22–31.

Wasserman, S. and Faust, K., 1994. *Social network analysis: Methods and applications*. Cambridge UP. 178-192.

Watts, D. and Strogatz, S., 1998. Collective dynamics of 'small-world'networks. *Nature*, 393 (6684), 440–442.

Webb, S., 2016. Horror traffic jam in Indonesia lasted 35 hours. https://www.mirror.co.uk/news/world-news/horror-traffic-jam-indonesia-lasted-8377678. Accessed: 23/02/2024.

Whitham, G., 1974. Linear and nonlinear waves. *Pure Appl. Math.*. 651.

World Economic Forum, 2016. The number of cars worldwide is set to double by 2040. https://www.weforum.org/agenda/2016/04/the-number-of-cars-worldwide-is-set-to-double-by-2040/. Accessed: 18/03/2024.

Wu, J., Gao, Z. and Sun, H., 2004. Simulation of traffic congestion with sir model. *Mod. Phys. Lett. B*, 18 (30), 1537–1542.

Wu, J., Gao, Z., Sun, H. and Huang, H., 2006. Congestion in different topologies of traffic networks. *Eur. Phys. Lett.*, 74 (3), 560.

Yildirimoglu, M. and Kim, J., 2018. Identification of communities in urban mobility networks using multi-layer graphs of network traffic. *Transp. Res. Part C: Emerg. Technol.*, 89, 254–267.

Youssef, M. and Scoglio, C., 2011. An individual-based approach to sir epidemics in contact networks. *Journal of Theoretical Biology*, 283 (1), 136–144.

Zang, J., Jiao, P., Liu, S., Zhang, X., Song, G. and Yu, L., 2023. Identifying traffic congestion patterns of urban road network based on traffic performance index. *Sustain.*, 15 (2), 948.

Zeng, C., Zhao, Z., Wen, C., Yang, J. and Lv, T., 2020. Effect of complex road networks on intensive land use in china's beijing-tianjin urban agglomeration. *Land*, 9 (12).

Zeng, G., Li, D., Guo, S., Gao, L., Gao, Z., Stanley, H. and Havlin, S., 2019. Switch between critical percolation modes in city traffic dynamics. *Proc. Natl. Acad. Sci.*, 116 (1), 23–28.

Zhang, Q., Karsai, M. and Vespignani, A., 2018. Link transmission centrality in large-scale social networks. *EPJ Data Science*, 7 (1), 33.

Zhang, Y. and Cheng, L., 2023. The role of transport infrastructure in economic growth: empirical evidence in the UK. *Transport Policy*. 223-233.

Zhang, Y., Li, S., Zhang, J., Ma, J. and An, H., 2022. Traffic-driven si epidemic spreading on scale-free networks. *Int. J. Mod. Phys. C*, 33 (08), 2250111.

Zhao, L., Li, W. and Zhu, Y., 2018. The robustness of worldwide metro networks. *J. Phys. Conf. Ser.*, 1113 (1), 012010.

Zheng, G., Chai, W., Duanmu, J. and Katos, V., 2023a. Hybrid deep learning models for traffic prediction in large-scale road networks. *Inf. Fusion*, 92, 93–114.

Zheng, J., Gao, Z. and Zhao, X., 2007. Modeling cascading failures in congested complex networks. *Phys. A: Stat. Mech. Appl.*, 385 (2), 700–706.

Zheng, S., Jiang, R., Tian, J., Li, X., Jia, B., Gao, Z. and Yu, S., 2023b. A comparison study on the growth pattern of traffic oscillations in car-following experiments. *Transp. B Transp. Dyn.*, 11 (1), 706–724.

Zheng, Z., Wang, Z., Zhu, L. and Jiang, H., 2020. Determinants of the congestion caused by a traffic accident in urban road networks. *Accid. Anal. Prev.*, 136, 105327.

Zhou, Y., Sheu, J. and Wang, J., 2017. Robustness assessment of urban road network with consideration of multiple hazard events. *Risk Anal.*, 37 (8), 1477–1494.

Zhu, D., Du, H., Sun, Y. and Cao, N., 2018. Research on path planning model based on short-term traffic flow prediction in intelligent transportation system. *Sensors*, 18 (12), 4275.

Zhu, J., Hu, L., Z.Chen and Xie, H., 2022. A queuing model for mixed traffic flows on highways considering fluctuations in traffic demand. *J. Adv. Transp.*, 4625690.